

1 CCGCAACCCC GACGGCGCCC CAAACGCTGT TCGCCGCGC GCCCGCCCA  
51 GCCCGGCCTC GCGCTGGTCC CGGTCTCGCC CCGCAGCCCT CGATCTCCCG  
101 TGACTTCCTC GGCCAGGCCG CCTGCGCCTC TGGGACCATG TTGCGCTGGC  
151 TGCGGGACTT CGCGCTGCCC ACCGCGGCCT GCCAGGACGC GGAGCAGCCG  
201 ACGCGCTACG AGACCTCTT CCAGGCACTG GACCGCAATG GGGACGGAGT  
251 GGTGGACATC GGCAGCTGC AGGAGGGGCT CAGGAACCTG GGCATCCCTC  
301 TGGGCCAGGA CGCCGAGGAG AAAATTTTCTA CTACTGGAGA TGTCAACAAA  
351 GATGGGAAGC TGGATTTTGA AGAATTTATG AAGTACCTTA AAGACCATGA  
401 GAAGAAAATG AAATTTGGCAT TTAAGAGTTT AGACAAAAT AATGATGGAA  
451 AAATTGAGGC TTCAGAAATT GTCCAGTCTC TCCAGACACT GGGTCTGACT  
501 ATTTCTGAAC AACAAAGCAG GTTGATTCTT CAAAGCATTG ATGTTGATGG  
551 GACAATGACA GTGGACTGGA ATGAATGGAG AGACTACTTC TTATTTAATC  
601 CTGTTACAGA CATTGAGGAA ATTATCCGTT TCTGGAAACA TTCTACAGGA  
651 ATTGACATAG GGGATAGCTT AACTATTCCA GATGAATTCA CGGAAGACGA  
701 AAAAAAATCC GGACAATGGT GGAGGCAGCT TTTGGCAGGA GGCATTGCTG  
751 GTGCTGTCTC TCGAACAAAGC ACTGCCCTT TGGACCGTCT GAAAATCATG  
801 ATGCAGGTTT ACGGTTCAAA ATCAGACAAA ATGAACATAT TTGGTGGCTT  
851 TCGACAGATG GTAAAGAAG GAGGTATCCG CTCGCTTTAG AGGGGAAATG  
901 GTACAAACGT CATCAAAATT GCTCCTGAGA CAGCTGTTAA ATTCTGGGCA  
951 TATGAACAGT ACAAGAAGTT ACTTACTGAA GAAGGACAAA AAATAGGAAC  
1001 ATTTGAGAGA TTTATTTCTG GTTCCATGGC TGGAGCAACT GCACAGACTT  
1051 TTATATATCC AATGGAGGTT ATGAAAACCA GGCTGGCTGT AGGCAAACT  
1101 GGGCAGTACT CTGGAATATA TGATTGTGCC AAGAAGATTT TGAAACATGA  
1151 AGGCTTGGGA GCTTTTACA AAGGCTATGT TCCCAATTTA TTAGGTATCA  
1201 TACCTTATGC AGGCATAGAT CTTGCTGTGT ATGAGCTCTT GAAGTCCTAT  
1251 TGGCTGGATA ATTTTGCAAA AGATTCTGTA AACCTGGAG TCATGGTGTT  
1301 GCTGGGATGC GGTGCCTTAT CCAGCACCTG TGGTCAGCTG GCCAGCTACC  
1351 CATTGGCTTT GGTGAGAACT CGCATGCAGG CTCAAGCCAT GTTAGAAGGT  
1401 TCCCCACAGC TGAATATGGT TGGCCTCTTT CGACGAATTA TTTCCAAAGA  
1451 AGGAATACCA GGACTTTACA GAGGCATCAC CCCAACTTC ATGAAGGTGC  
1501 TCCCTGCTGT AGGCATCAGT TATGTGGTTT ATGAAAATAT GAAGCAAAT  
1551 TTAGGAGTAA CCCAGAAATG ATGTTGCATT TTTTGCTTTA GCCTGATAAT  
1601 TGAAACTTTC AACAACTCTC GGAGTGACTT TTTCTCCTCG AATTGAAACA  
1651 AGTCTATGGC AAAAGAAGCT GCATTTTTTT CACAAAAGG AAGACGGTAA  
1701 CAATGTCAC TTCAAACCTT TGGGCTAAAT TATATGTACA CAGAAATGTT  
1751 CAAATCATA GTTTTAATGT GTTTTGAAAA GGCCACACAA TTATACTTTA  
1801 TCTTTTCTTA ATAATCCTGC AAATCTCTGC CCTGAATCCG AAATCTGAAA  
1851 ATGTACTGGC TTGAACAAA TTTGTTTGT GTGTAGAGT TATAAATCAT  
1901 TAATCTTTAT TTCGGGTGGT TTACGTTTAT GCCAGTTCCT TTATATTTAA  
1951 ATTTCTTGTT TTATATATTT TGAATGTCTT TATAGATTTT TTTAAATTTT  
2001 CTTATAGAAC CATTAATAGA AAATCATTAC ATTTAAATA TACCTTACAG  
2051 CAAAAGCATC CAAATAAGTA TAGGGTTTAT GTCCTTATTT TTCTTTCAGC  
2101 TGAATACGAA TGAACACAGT GGTGGAATTT CTGAAGGGAA GTGATGAAAT  
2151 TATATTTATT TCAGTGGGCA CTTTTCATT TTACCACTGT ACCATTATTT  
2201 GGTTCCTGGA GTTATACACT AATTTTCAGT ATATTACTGT TAAATTACCA  
2251 ACACAAGGCA ATTTATTTGA AAGATTCCGT TTATCCTGCC ATTGCTTTGA  
2301 AAAGCAGCAG GAAACGAAAT TTTTGTACTT GTATCAGCTT CTGCAGAGCA  
2351 TCTTTGTTTT CTTTGTCTCT TTGTTTCTTA CTTTGTGAAT CAGATTCCGT  
2401 TTTAGTCAGG AAGACTTCTT GGGACCATTC TTAGTAACCT GAAATTTCTT  
2451 TTTTAATTGC ATGAAGTGGA TTGATCATGA GCAAGTGATG GGCTTTATTT  
2501 CTCCCTCACT GGTGAATATC CTTTGAACCT GCTGTTTGCA ATATGGGCAG  
2551 CCACAAAGGG GGAGAGATGC CTATTAAATC GGCGGGGTGT ATGACTTCTG  
2601 AAAACATTGG ATACCCTATT TTGAAAAGGG AAAGGCCCAA TTTGGGAAA  
2651 CATATACCAA TGCATGATTT CTG

# FEATURES:

5'UTR: 1-137  
Start Codon: 138  
Stop Codon: 1569  
3'UTR: 1572

# **HOMOLOGOUS PROTEINS:**

## Top BLAST Hits:

	Score	E
CRA 335001098641184 /altid=gi 11360341 /def=pir  T50686 peroxis...	927	0.0
CRA 11000479457833 /altid=gi 6841066 /def=gb AAF28888.1 AF12330...	834	0.0
CRA 18000005183605 /altid=gi 7504235 /def=pir  T22688 hypotheti...	432	e-120
CRA 1000682325160 /altid=gi 7499323 /def=pir  T21074 hypothetical...	377	e-103
CRA 89000000196990 /altid=gi 7294582 /def=gb AAF49922.1  (AE003...	348	9e-95
CRA 150000075553401 /altid=gi 9758252 /def=dbj BAB08751.1  (AB0...	339	5e-92
CRA 335001098657884 /altid=gi 11358611 /def=pir  T49871 peroxis...	330	2e-89
CRA 163000046661776 /altid=gi 10176874 /def=dbj BAB10081.1  (AB...	326	4e-88
CRA 105000014652720 /altid=gi 10798831 /def=dbj BAB16462.1  (AP...	200	3e-50
CRA 335001098655048 /altid=gi 11277065 /def=pir  T47703 Ca-depe...	199	6e-50

## BLAST dbEST hits:

gi 10145202 /dataset=dbest /taxon=96...	1108	0.0
gi 1437155 /dataset=dbest /taxon=9606 ...	801	0.0
gi 10333851 /dataset=dbest /taxon=96...	745	0.0
gi 8469752 /dataset=dbest /taxon=960...	363	8e-98
gi 11684041 /dataset=dbest /taxon=96...	307	4e-81

## **EXPRESSION INFORMATION FOR MODULATORY USE:**

library source:

### Expression information from BLAST dbEST hits:

gi|10145202 Placenta Choriocarcinoma  
gi|1437155 Retina  
gi|10333851 Uterus leiomyosarcoma  
gi|8469752 Breast  
gi|11684041 Ovary fibrotheoma

### Expression information from PCR-based tissue screening panels:

Leukocyte

```

1 MLRWLRDFAL PTAACQDAEQ PTRYETLFQA LDRNGDGVVD IGELQEGLRN
51 LGIPLGQDAE EKIFTTGDVN KDGKLDFEFF MKYLDHEKK MKLAFKSLDK
101 NNDGKIEASE IVQSLQTLGL TISEQQAELI LQSIDVDGTM TVDWNEWRDY
151 FLFPVTDIE EIIRFWKHST GIDIGDSLTI PDEFTEDKK SGQWWRQLLA
201 GGIAGAVSRT STAPLDRKI MMQVHGSKSD KMNIFGGFRQ MVKEGGIRSL
251 WRGNGTNVIK IAPETAVKFW AYEQYKLLT EEGQKIGTFE RFISGSMAGA
301 TAQTFIYPME VMKTRLAVGK TGQYSGIYDC AKKILKHEGL GAFYKGYVPN
351 LLGIIPYAGI DLAVYELLKS YWLDNFAKDS VNPGVMVLLG CGALSSTCGQ
401 LASYPLALVR TRMQAQAMLE GSPQLNMVGL FRRIISKEGI PGLYRGITPN
451 FMKVLPAVGI SYVVYENMKQ TLGVTQK

```

**FEATURES:**

**Functional domains and key regions:**

[1] PDOC00001 PS00001 ASN\_GLYCOSYLATION  
N-glycosylation site

254-257 NGTN

[2] PDOC00005 PS00005 PKC\_PHOSPHO\_SITE  
Protein kinase C phosphorylation site

Number of matches: 2

1	229-231	SDK
2	475-477	TQK

[3] PDOC00006 PS00006 CK2\_PHOSPHO\_SITE  
Casein kinase II phosphorylation site

Number of matches: 8

1	22-25	TRYE
2	65-68	TTGD
3	121-124	TISE
4	157-160	TDIE
5	170-173	TGID
6	179-182	TIPD
7	185-188	TEDE
8	227-230	SKSD

[4] PDOC00008 PS00008 MYRISTYL  
N-myristoylation site

Number of matches: 16

1	52-57	GIPLGQ
2	119-124	GLTISE
3	171-176	GIDIGD
4	201-206	GGIAGA
5	202-207	GIAGAV
6	245-250	GGIRSL
7	253-258	GNGTNV
8	283-288	GQKIGT
9	295-300	GSMAGA
10	322-327	GQYSGI
11	326-331	GIYDCA
12	359-364	GIDLAV
13	392-397	GALSST
14	399-404	GQLASY
15	442-447	GLYRGI
16	446-451	GITPNF

[5] PDOC00018 PS00018 EF\_HAND  
EF-hand calcium-binding domain

Number of matches: 3

1	32-44	DRNGDGVVDIGEL
2	68-80	DVNKDGKLDFFEF
3	99-111	DKNNDGKIEASEI

**Membrane spanning structure and domains:**

Helix	Begin	End	Score	Certainty
1	292	312	1.053	Certain
2	345	365	0.613	Putative
3	381	401	1.544	Certain
4	446	466	0.733	Putative

**BLAST Alignment to Top Hit:**

>CRA|335001098641184 /altid=gi|11360341 /def=pir||T50686 peroxisomal  
Ca-dependent solute carrier [imported] - rabbit  
/org=rabbit /taxon=9986 /dataset=nraa /length=475  
Length = 475

Score = 927 bits (2371), Expect = 0.0

Identities = 454/477 (95%), Positives = 466/477 (97%), Gaps = 2/477 (0%)

Query: 1 MLRWLRDFALPTAACQDAEQPTRYETLFQALDRNGDGVVDIGELQEGLRNLGIPLGQDAE 60  
MLRWLR F LPTAACQ AE PTRYETLFQALDRNGDGVVDI ELQEGL++LGIPLGQDAE  
Sbjct: 1 MLRWLRGFVLPTAACQGAEPTRYETLFQALDRNGDGVVDIRELQEGLKSLGIPLGQDAE 60

Query: 61 EKIFTTGDVNKGKLDFFEEFMKYLDHEKKMKLAFKSLDKNNDGKIEASEIVQSLQTLGL 120  
EKIFTTGDVNKGKLDFFEEFMKYLDHEKKMKLAFKSLDKNNDGKIEASEIVQSLQTLGL  
Sbjct: 61 EKIFTTGDVNKGKLDFFEEFMKYLDHEKKMKLAFKSLDKNNDGKIEASEIVQSLQTLGL 120

Query: 121 TISEQQAELILQSIDVDGTMTVDWNEWRDYFLNPNVTDIEEIIREFWKHSTGIDIGDSLTI 180  
TISEQQAELILQSID DGTMTVDWNEWRDYFLNPNV DIEEIIREFWKHSTGIDIGDSLTI  
Sbjct: 121 TISEQQAELILQSIDADGTMTVDWNEWRDYFLNPNVADIEEIIREFWKHSTGIDIGDSLTI 180

Query: 181 PDEFTEDKKSGQWWRQLLAGGIAGAVSRTSTAPLDRKIMMQVHGSKSDKMNIFFGGFRQ 240  
PDEFTE+E+KSGQWWRQLLAGGIAGAVSRTSTAPLDRK+MMQVHGSKS MNIFGGFRQ  
Sbjct: 181 PDEFTEEERKSGQWWRQLLAGGIAGAVSRTSTAPLDRKVMQVHGSKS--MNIFGGFRQ 238

Query: 241 MVKEGGIRSLWRGNGTNVIKIAPETAVKFWAYEQYKLLTEEGQKIGTFERFISGSMAGA 300  
M+KEGG+RSLWRGNGTNVIKIAPETAVKFW YEYKLLTEEGQKIGTFERFISGSMAGA  
Sbjct: 239 MIKEGGVRSRLWRGNGTNVIKIAPETAVKFWVYEQYKLLTEEGQKIGTFERFISGSMAGA 298

Query: 301 TAQTFIYPMEVMKTRLAVGKTGQYSGIYDCAKKILKHEGLGAFYKGYVPNLLGIIPYAGI 360  
TAQTFIYPMEVMKTRLAVGKTGQYSGIYDCAKKILK+EG GAFYKGYVPNLLGIIPYAGI  
Sbjct: 299 TAQTFIYPMEVMKTRLAVGKTGQYSGIYDCAKKILKYEGFGAFYKGYVPNLLGIIPYAGI 358

Query: 361 DLAVYELLKSYWLDNFAKDSVNPVGMVLLGCGALSSTCGQLASYPLALVRTRMQAAMLE 420  
DLAVYELLKS+WLDNFAKDSVNPGV+VLLGCGALSSTCGQLASYPLALVRTRMQAAMLE  
Sbjct: 359 DLAVYELLKSHWLDNFAKDSVNPGLVLLGCGALSSTCGQLASYPLALVRTRMQAAMLE 418

Query: 421 GSPQLNMVGLFRRIISKEGIPGLYRGITPNFMKVLPAVGISYVVYENMKQTLGVTQK 477  
G+PQLNMVGLFRRIISKEG+PGLYRGITPNFMKVLPAVGISYVVYENMKQTLGVTQK  
Sbjct: 419 GAPQLNMVGLFRRIISKEGLPGLYRGITPNFMKVLPAVGISYVVYENMKQTLGVTQK 475

>CRA|11000479457833 /altid=gi|6841066 /def=gb|AAF28888.1|AF123303\_1  
(AF123303) calcium-binding transporter [Homo sapiens]  
/org=Homo sapiens /taxon=9606 /dataset=nraa /length=411  
Length = 411

Score = 834 bits (2132), Expect = 0.0

Identities = 409/410 (99%), Positives = 409/410 (99%)

Query: 8 FALPTAACQDAEQPTRYETLFQALDRNGDGVVDIGELQEGLRNLGIPLGQDAE EKI FTTG 67  
F LPTAACQDAEQPTRYETLFQALDRNGDGVVDIGELQEGLRNLGIPLGQDAE EKI FTTG  
Sbjct: 1 FVLPTAACQDAEQPTRYETLFQALDRNGDGVVDIGELQEGLRNLGIPLGQDAE EKI FTTG 60

Query: 68 DVNKGKLDFFEEFMKYLDHEKKMKLAFKSLDKNNDGKIEASEIVQSLQTLGLTISEQQA 127  
DVNKGKLDFFEEFMKYLDHEKKMKLAFKSLDKNNDGKIEASEIVQSLQTLGLTISEQQA  
Sbjct: 61 DVNKGKLDFFEEFMKYLDHEKKMKLAFKSLDKNNDGKIEASEIVQSLQTLGLTISEQQA 120

Query: 128 ELILQSIDVDGTMTVDWNEWRDYFLNPNVTDIEEIIREFWKHSTGIDIGDSLTI PDEFTED 187  
ELILQSIDVDGTMTVDWNEWRDYFLNPNVTDIEEIIREFWKHSTGIDIGDSLTI PDEFTED  
Sbjct: 121 ELILQSIDVDGTMTVDWNEWRDYFLNPNVTDIEEIIREFWKHSTGIDIGDSLTI PDEFTED 180

Query: 188 EKKSQWWRQLLAGGIAGAVSRTSTAPLDRKIMMQVHGSKSDKMNIFFGGFRQ MVKEGGI 247

EKKSGQWWRQLLAGGIAGAVSRTSTAPLDRCLKIMMQVHGSKSDKMNIFFGFRQMVKEGGI  
 Sbjct: 181 EKKSGQWWRQLLAGGIAGAVSRTSTAPLDRCLKIMMQVHGSKSDKMNIFFGFRQMVKEGGI 240  
 Query: 248 RSLWRGNGTNIKIAPIETAVKFWAYEQYKKLLTEEGQKIGTFERFISGSMAGATAQTFFIY 307  
 RSLWRGNGTNIKIAPIETAVKFWAYEQYKKLLTEEGQKIGTFERFISGSMAGATAQTFFIY  
 Sbjct: 241 RSLWRGNGTNIKIAPIETAVKFWAYEQYKKLLTEEGQKIGTFERFISGSMAGATAQTFFIY 300  
 Query: 308 PMEVMKTRLAVGKTGQYSGIYDCAKKILKHEGLGAFYKGYVPNLLGIIPYAGIDLAVYEL 367  
 PMEVMKTRLAVGKTGQYSGIYDCAKKILKHEGLGAFYKGYVPNLLGIIPYAGIDLAVYEL  
 Sbjct: 301 PMEVMKTRLAVGKTGQYSGIYDCAKKILKHEGLGAFYKGYVPNLLGIIPYAGIDLAVYEL 360  
 Query: 368 LKSYWLDNFAKDSVNPVGMVLLGCGALSSTCGQLASYPLALVRTRMQAQA 417  
 LKSYWLDNFAKDSVNPVGMVLLGCGALSSTCGQLASYPLALVRTRMQAQA  
 Sbjct: 361 LKSYWLDNFAKDSVNPVGMVLLGCGALSSTCGQLASYPLALVRTRMQAQA 410

Score = 80.0 bits (194), Expect = 6e-14  
 Identities = 80/388 (20%), Positives = 156/388 (39%), Gaps = 59/388 (15%)

Query: 95 FKSLDKNNDGKIEASEIVQSLQTLGLTISEQQAELILQSIDV--DGTMTVDWNEWDRDYFL 152  
 F++LD+N DG ++ E+ + L+ LG+ + + E I + DV DG +  
 Sbjct: 21 FQALDRNGDGVVDIGELQEGLRNLGIPLGQDAEEKIFTTGDVKNKDGKL----- 68  
 Query: 153 FNPVTDIEEIIIRFWKHSTGIDIGDSLTPDEFTDEKKSGQWWRQLLAGGIAGAVSRTST 212  
 D EE +++ K + EKK ++ L +  
 Sbjct: 69 -----DFEEFMKYLK-----DHEKKMKLAFKSLDKNNDGKIEASEIV 105  
 Query: 213 APLDRCLKIMMQVHGSKSDKMNIFFGFRQMVKEGGIRSLWRGNGTNIKIAPIETAVKFWAY 272  
 L L + + ++ +I V R + N I E ++FW +  
 Sbjct: 106 QSLQTLGLTISEQQAELILQSIDVDGTMTVDWNEWDRDYFLFNPVTDI----EEIIRFWKH 161  
 Query: 273 EQYKKL-----LTEEGQKIGTFER-FISGSMAGATAQTFFIYPMEVMKTRLAV-GKT 321  
 + TE+ +K G + R ++G +AGA ++T P++ +K + V G  
 Sbjct: 162 STGIDIGDSLTPDEFTDEKKSGQWWRQLLAGGIAGAVSRTSTAPLDRCLKIMMQVHGSK 221  
 Query: 322 GQYSGIYDCAKKILKHEGLGAFYKGYVPNLLGIIPYAGIDLAVYELLKSYWLDNFAKDSV 381  
 I+ ++++K G+ + ++G N++ I P + YE K ++  
 Sbjct: 222 SDKMNIFFGFRQMVKEGGIRSLWRGNGTNIKIAPIETAVKFWAYEQYKKL----LTEEGQ 277  
 Query: 382 NPGVMVLLGCGALSSTCGQLASYPLALVRTRMQAAMLEGSPQLNMVGLFRRIISKEGIP 441  
 G G+++ Q YP+ +++TR+ A+ + + ++I+ EG+  
 Sbjct: 278 KIGTFERFISGSMAGATAQTFFIYPMEVMKTRL---AVGKTGQYSGIYDCAKKILKHEGLG 334  
 Query: 442 GLYRGITPNFMKVLPAVGISYVYENMK 469  
 Y+G PN + ++P GI VYE +K  
 Sbjct: 335 AFYKGYVPNLLGIIPYAGIDLAVYELLK 362

#### Hmmer search results (Pfam):

Model	Description	Score	E-value	N
PF00153	Mitochondrial carrier proteins	305.4	3e-88	1
PF00036	EF hand	50.7	1.7e-12	3
PF00404	Dockerin domain type I	9.7	0.26	1
PF01978	Protein of unknown function	2.7	9.5	1

#### Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00036	1/3	27	51 ..	5	29 .]	18.7	0.002
PF00404	1/1	67	85 ..	1	22 [.]	9.7	0.26
PF00036	2/3	61	87 ..	3	29 .]	19.7	0.001
PF00036	3/3	90	118 ..	1	29 [.]	17.2	0.0051
PF01978	1/1	110	121 ..	1	13 [.]	2.7	9.5
PF00153	1/1	193	472 ..	1	313 [.]	305.4	3e-88

1 AACCCATGTT AGTGTGCAGT TCTGCTGGCA CACACATGCA GTTGTGTAAC  
51 CACTACCACC AAAAGCAAGA TGTAATAATAG CTCCATCACC CCCACAAGCC  
101 TTCTGATGCT CTTTTGTCTAT CAATTCCTT CCCGCTAGTC ACAACTGGTA  
151 ACTACTGATT TGTTTTCTGT CCCTATAGTT TTGCCTTTTC CAGAATGTCA  
201 TTGTTGACAG GTATCAGTAA TTCATTCCCTT TTTATTGCTA ATTACTATCT  
251 CACTGTATGA ATGCAACACA GGTTGTTTAC CAGTTCACCC GTTAAAGAAC  
301 ATTTTGTTC TGCGCTTGAC AGTTATGAAT AGAACTGCTA TAAACCCTCA  
351 AGTAAAAGTT TTGGTGTGAA GATAATTTTC TCAGCAAAAA CGCTGACAGG  
401 TAATTTTTCT AAGTATTACT TTTTAAAAA AGTAAAATAG CCTGTAGCCC  
451 CAGCTACTCA GGAGGCTGAG GCAGGAGAAT AGCTTGAACC CAGGAGGCGG  
501 AGGTTGCAGT GAGTTGAGAT TGTGCCACTG CATTCCAGCC TGGGCGACAG  
551 AGTCTAGTG TCTCAAGAA AAAAAAATA AATAACAAAT AAATAAAAAG  
601 TAAAATGAAA GCATGTAAAGT GTAAGATGAC TAGTTCACGC AACCTCTCTT  
651 CAAGTACAGA GTATTCTAGAG TAGAGATTAA AAGAGGTTTT CAAGGACAGA  
701 GAAAATTTGA AGTTTGAAGG CAGTTCCAAA GGAAGGCAAT GATTCTTAAT  
751 AAGACTGGAA GTTGGAAAGTA ATATAAAAAAG ATAAATCAGT TTCAAGATGA  
801 TTTTACTAAG CAGGCAGCCC TTAATTTACA AATTCTAGAT TCATACATAT  
851 CTTAAACATA CAAAATGATA TGAGGAGAGG TAAGTTCAGG GTCTGAGTTC  
901 CTGGCTGTTG TTGGAACTGA TTTCTGTGTA GTGATTGAGA AGATGTGAGA  
951 CACCCTAATT TACAAGTACA GAGGTATCTT CTTTTCTGCA AACAGCAGTA  
1001 CAACAATAGT TCCTCTTACG CAGCTGTGAA TGAACAGGAT TATTACAATT  
1051 AATGATATCT CATTGTATTG GCGCCTTAGA GAATTAAGAC CTTTCACACC  
1101 TAATATACAA CTTTGTGTG AAGGCAGATA TTTATATTCT CATTTTACTG  
1151 ATGAGAGACT ACCCGGAGAC GCTATGTCAC ACCTGAAGGA TTAGGTACTT  
1201 TCTCTGTTAA GTCCAATGTT CCTTCCGTTA TTCCATGCTA GGCAGTAATA  
1251 AGTTCTGTCT TGCCTGAGTA ATAAGCTCCA AACCTCGGAA CTGCACCCAT  
1301 CTTGAGAAGG AGGAGGGCGC TGTGGTTTTT TCTGATAAGT GCAGCTGGCA  
1351 GACACTCTAT ACGCTTAATC ACGGGCAAAT CCTACCTAAG CTGCCTACCA  
1401 AACTAGTCTT TCTTTTCCCC GTTGCCACG CAGATGGCTG TTGATCTTTT  
1451 CTGCAACAAA TCCAGGAGTT TCTCCTTTTT GTTTTATAAT TGCTCCAATA  
1501 GATGCTTTAG GATTTAATC TCTGCTTTTT AAAGCAGAAAT CGCCATCCCA  
1551 GGTGTGCAAC CACGAAAAA TTAGACATCC GTGAGAGACA ATGCCCTCCA  
1601 TGGCCAGTT TCCAGGCAGA GAGAAGCAGC TCTGGGCTGA CCGCCAAGGC  
1651 TCCGGCCCGA GAGGCTCTTT AAGTGGAGTA ACCAGTCTTC AAGACCCCGC  
1701 TCCAAGCCA CCGACGCGCT GACGCTGCAG CCCTGGACCT GCTGGGGGCC  
1751 TCTTCCTCGG ACCCGCATGC TGACAGCGGG ACTGGCAACT GGGCAGAGGT  
1801 CGACCCCGGG TCCGCACAGC ACCTCCCGAG ACCCAGCTCC CAGCTCCCTC  
1851 ACTTCCGGCT CTCTGGAGG GGGCCCGGCC AGTGCCGCGG AGGCCAGCGC  
1901 GCGGAGCTCC TCCCGAGCAG CGGCGGGACG GCCACACCTT GCGCGCCGCG  
1951 CGGGCTCGGG TGGGGTCTCC GCTCCTGCGC CCTGCGCGCC GCAGCCGCAC  
2001 CCCCAGCGGC GCCCCAAACG CTGTTGCGCC GCGCGCCCG CCCAGCCCGG  
2051 CCTCGCGCTG GTCCCGGTCT CGCCCCGAG CCCTCGATCT CCCGTGACTT  
2101 CCTCGGCCAG GCCGCTGCG CCTCTGGGAC CATGTTGCGC TGGCTGCGGG  
2151 ACTTCGTGCT GCCACCGCG GCCTGCCAGG ACGCGGAGCA GCCGACGCGC  
2201 TACGAGACCC TCTTCAGGC ACTGGACCGC AATGGGGAGC GAGTGGTGGA  
2251 CATCGCGCAG CTGCAGGAGG GGCTCAGGAA CCTGGGCATC CCTCTGGGCC  
2301 AGGACGCCGA GGAGGTGGGT CGCCGCCGGG GCGCCGCTG AGCGTAGGGA  
2351 GGGCTGCGGG CGCTGGGGAC ACTGCGAGGA CCGAGGAGGG CGGCGGCTTG  
2401 AGGCGTTGCC AGGAGAGGAA GGAGGAACTG TGGCGCCAG CGCTCCGGTG  
2451 GCTTCAGAAA CTCGGGCGTG GGGCCGCGAC CGGCGACCCC GGTAACAGAA  
2501 GTGGGTGATA ATACGAAAGT CTAATGGTAT TTGTCCAGAT AAAATGAGTG  
2551 TTGTGGACAC TCTGGCCAC GGGCACTGTT AAATTTTAA GACACTTTTG  
2601 TCCTGAATCC ATCCCAGGTT CTTTGTTCCT TGTTTTAATA CCTTGCAGAC  
2651 ATGTAATCCG TTTTAGCTGT CAGACTTCAG TGGGTCCCAA GTTTTGTATA  
2701 AAGGCGCACA CATTGATCT CTTTCGAAGC TGCTTTGTTA CAGCAGCTAT  
2751 GTGTATTGTC TACTGTTTGA AAATGTTTG AAAACCAATC GCGTGTTCCT  
2801 CCCACTTCCT GTTGAGAAGG AATGGCGGCA TTCCATTGTT TAAGACATTC  
2851 CTAGGTTAAT GCCCTAGGTA CATAAATTGA TCTGAAGGGT TGAATGAGC  
2901 TGCGACTGAG CAATTTCAAT TTCTCTGAGT CATCTTAAT GTGCCCTGA  
2951 ACTTCTGCCC CTTTAGTAGG GTGGAGATAT GTGGAACCTC TCCAACCTG  
3001 TTGAAGCGTT CCCTGACACT GGCATTCTCT TATCCAAAGA GGGAAAGTGA  
3051 TTAGGTTACT ATGAGGGCCA ACAACTGTTA TATAGTTATA TTTCATTCT  
3101 CTTTAAATGT CTTTGGTAGT TATAGGCTC TTCAGTTTAC TGTTTCTTCT

FIGURE 3, page 1 of 42

3151 AGAGTCAGAT TTAGTAAGTT ACAATTTTTT TTGAAACTGC CTGTTCTGTC  
3201 CAAGGTTTCAT AATACTCACC GATGATTTTA TAACACTTCT GACTGAATCT  
3251 GTAGGTAGGT TCTCTATTTC ATTCCTCATA TCTATCCTTT TCTCCCCCTC  
3301 AATCTTGCCA AAGTTTTGTG TATTTTATTC ATACTTTGAA GGAACCAACT  
3351 TTTGGTACTT TGTGCTGATT GTCCCAGAAA TGGCCAGTT GGAGTTCCCC  
3401 ACCATGTCCA ATCATTGGCT GGAAGCAGCC CAGGAAAGGG ACGACCTTGC  
3451 TGCAGTGCAT CAGCAGATGC CAGGGTTAGA GGCTAGAGAG TGGAAGTCAA  
3501 CTGTGTTTCT CACAGTAGGT GCCTTTGAAG GGAGATCTCA GTGGTACAAC  
3551 TCCATGGTCC CTACAATATA CAAAAGCTCT TTGGAGTGCT CAATGATTTT  
3601 TAAGATTGTA AAGGGATCCT GAGATCAAAA AGCTTGAGAA TTGCTGCTGT  
3651 ATCACCATT TACGTAAC GCATCATATT CTGTTATATG TTTGTGTCAT  
3701 AGTAGTAGTT ACCAATTCTT TTTAAATCAC CTTTACTTTT ATTGATAGTT  
3751 TAAAAACGAT TGTAAGTGAA ATTGCAATGG ATGTCCTTTG TATTCATTTT  
3801 CTCATTCTGG TCCAGTTACT TTCGTAGGAT AAATTTTGAG GAGTGGACAT  
3851 TGCTGAGTCT GAAGGTAACA CACATTTTAA ACTGGGATAC GTATTGCCCT  
3901 TCGGAAACCT TAGACCCATT TTCACTCTTT TGACTGACAG TGCTTGCTTC  
3951 TCCACATCCT CGCTCATFCA GGGTATCAGT CTTTGTAAG TCTCCTATTC  
4001 TGCAGGTGAA ATTCCTTTTC ATTTCTGCT TTAGTCCATT TAGTGTGCT  
4051 ATAGTGGAA ATCTGAGACA GGGTAATTTA TAAAGAAAAG ACATTTATTT  
4101 AGCTCACAGT TCCGCAGGCT GGGAAGTTTA AGAAGCGTGG TGCTGGCATC  
4151 TGCTGGACTC CTGGGGAGGG CTTTCTGCT GTGTCACAAC ATGGTGGA  
4201 GTCAAAGTGG AAGTGGACAT GTGTGAAGAA GCAAATCCG AGGGGTGTCC  
4251 TGGCTTTATA GCAACCCAGC CTCGAGGGAA CTGATCCATT ACTGAGGGAA  
4301 CTAATTCAGT CTCATGAGAG AGAGAACTCA CTCACTACTG CAAGAATGAC  
4351 ACCAAGCCAT TCATGAGGGA TCTGCCTCCG TAACCCTGAC ACCTCCTGCT  
4401 AGGTCCCTCC TCCCAACACG GCCACATCAG GGATCAGACT TCAACATGAG  
4451 TTTTGTGGG GACAAACAAA ACGTAGCACT TGCTTTGCCT TTTGGTTCTA  
4501 TTCACATCCT CCACAGGATT GCATTATGCC TACCCATTTG GTGAGGGCAG  
4551 TCTTCTTTAA TTGGTTTACT GATTCAAATG CTACCTCCT CCAGAGACAT  
4601 CCTCACAGAC ACACCCAGAA ATCATGTTTT ACCAGTTATC TGGGCATCCC  
4651 TTAGTCCAGA CGAGTTGATA CATAAAATTA ACCATCACAC ATGGGATAGA  
4701 ATTAGGATTA CACAGTCAAC CTTTATGGGA GAAAATTTCA GAGGCATGTC  
4751 AGGGGTTTAT GTAATGTCAA GGAGTGAGGA CATTGGCTAC TTGAGCATAG  
4801 AAATGAGAAC TGTGGGGTGA CTCTTCGGTG GAAAGTTTCA AGGTAGTAGT  
4851 TTGTATCTAA GCCAAATACT CAGCTTGAAG CAAAATCTCT ATAAATTTTC  
4901 ATCTGATTG ATCTCATCTC CGTGTTTCCA AGCATTTGTA ATGAATTGAG  
4951 CATTTAGAAG AGAACAAATT TCTGTTTAA TTTCTTTAGA TTTTAGATGG  
5001 AAAGAATGTA GAAATAAGAG TAGAATGTAG AAATAGGTAT AAAGAATATA  
5051 ATAGGTAACC ATTACTAAGT GTTCCAGAAT TATCCAGGGA AGAGAAAAGA  
5101 ATTCAAGGCA AGTCCTGAGA CAAAATTAAG AACCAATTGG AAGTGAAAGC  
5151 GCTACATTTT TTTTTTCTGG TATGACCTTT CTTTTCTATA TGTTCCAAAT  
5201 CTCCTCATA TGAAATTAGT GAAAAATTAA AGTTAAAAAT TAGAGAAAAT  
5251 TCACATTAAG TTCTCCTAGG ACTCAGTAGT ATAAGGGTAT AGACTGAGAG  
5301 TAGAATGTAG TGTGAGAAAC AGGAGATACA GTATTTAACC ATTACTAATT  
5351 CTCTTATACT TGTCTAGTAA TCCTATTTC TTTTAAAGT CTTCACTTAT  
5401 TTTCTCTTTA CGCACCTCCT TCTCCCTCTT GTCTTCCTCC TTCTACCCCC  
5451 ATCTTTCTTC CTGTGGAGCC TTCATGAATG GGATTAGTGC TTGTATAAAA  
5501 GTGACCTGGA AGACCTTCCT TGCCCCTTCC ACCATGTGAG GACACAGTGA  
5551 GAAACAGTG GTCCATGGAA CCGGAAAGTG GGTCCCTACT AGACAGTAAA  
5601 TCTCCTAGCA CTTGATCTA GGACTTCCAG TGTCTGGAAC TGCAAGAAAT  
5651 CAATGCTTAT TGTTTAAGTA AGCCAGTAGT ATTTTGTCA TAGCAGCCCA  
5701 GTTGGACTAG GACAATTACC AAGAGCAAGA AGGGAAGCAG CAAGCTACAA  
5751 GAGAGTCCG TCCTTGGTGT AAATTGACCG TGTAATCCTT GTCAAGTTTG  
5801 AGCCTTACTG GAGCTTTACT TTCTTATTCT TAAAATGCAG ATATCTTGCC  
5851 TGCATCCTGG ACAGAGCTTT TAACAAGGTC ATATGTTGCA GAATATGAAA  
5901 GTTCATGTTA AAAAACCTT TAAAATGTGG TATCCCATTT ACTAGCTGGT  
5951 GAACTTCTTG AGGAACCTCT GTGCCATGG GTATGAAGTG TATGCTGAAT  
6001 GATCACCCAA TGTTAGAGGA GTGGGTGGAC TGGTAACCTG ATTTAAGGGC  
6051 CATTCTAACT CTTACATTCT ATGATTTTTT TAATTCTGTC TTTAAGTTTT  
6101 TACATTTACA ATCACAGAAA AAATAGTCAC ATAGAAGAAT AGTAGCTTAG  
6151 CAAATGTTTA TTGCATTGAG TGGAATCAGG ATTTCACTCC ATTAAGTAAT  
6201 TCCTCTGTTA ACAAAGAGGG TTCATTTCAT TTTTATTTCA TTAATATTGC  
6251 TTTTTTTTTT TTTTTTCTGG AGACAGAATC TTGCTCTATC ACCAAGGCTG

FIGURE 3, page 2 of 42



6301 GAGTGCAGTG GTGCGATCTC GGCTCACTGC AGCCTCTGCT TCCTGGATTG  
6351 AAGCGATTCT TGTGCCTCAG CCTCCCAGC AGCTGAGATT ACAGGCACAT  
6401 GCCACCACAC CTGGTTAACT TTTGTATTTT CTAGTAGAGA TGGGATTTTG  
6451 CCATGTTGGT CAGGCTGGTC TTGAATTCCT GGCCTCTAGT GATCTGCCTG  
6501 CCTCTGCCCT TGAAAGTGCT AAGATTACAG GCATGAGCTA CCATGGCCAG  
6551 CCCATTTCCCT TAATATTTTA ATTGTCAGAC ATGTTATGGT TTCTGGCACA  
6601 ATATTAAGAA GACATGATAT GAAATCACAG GGTGAATTTT AGGGCATCAC  
6651 AACAGAAAGA TTATGGTATA AGAAAAACAA TGGAATTTCA ACTACATTTT  
6701 TGTCAAATGT TCTAAAATAT ATAAAATCTG TATCTTTTGT GTTCTCTCCT  
6751 GATTTATATT CTAAATTTGA TGTTATCCTT CTCTGCAGAA ATAAAGTGTC  
6801 TGAAAGAAATG AAAAAATGG AAGAATCTT TAGTAAGGTA TAAAAACCC  
6851 TTTCTATCTT TTTAGCATTG TAAGCCTTTT GTCACCTTTT CAAACTCCCA  
6901 ACATGCCATA TTCCCTGACT AGGCCACAGC CATGTACATT GATCCCTTTA  
6951 TTTTCTTCTC TCTGCCTGAG ATTTCTCTCA TTCCCCCTTC TCTGCCTGGT  
7001 ATATGATTGC CCATTGTTTA AGGCCCCAAC TCACCTTTAT AATCTTCCTA  
7051 GCCCACTTTC TTTATCGGTA TTCCAGAAAA AACAAAAGAA GCTTCCACAA  
7101 GACAACATTC TGTAATACAC TGCTTAACTT CTTTGGACCC TGCTGAGTTC  
7151 AAAAATCTTA TCTTTTAAAG GATTGAATGG AGTCCACCAA GGTATCTATA  
7201 TTTGACAGGA TTTATGAAAA CAAAAGGATT TGTTGAGAAA GTTTGAAGCC  
7251 TAACTCTGAA ACGTGGATCA TAGTGTTTAC TACACATTAA CTGTTTtagt  
7301 GGATGTAATA GTTATTATTA TAGGCTGTGG AATCAGAACA GGGTTCAAAT  
7351 GTTTTCACCG CTTGCTAGAC TGTGGCCTTG GGCATGTTAT TTAATGCCTG  
7401 GAGGCCTCAA ATGTTAACTA GGAATGGTAA GACCTACCCA GTAACCTAGC  
7451 ATAAATAGTA AATTCAATCA TTTAATGTTT TCAAACAGTG CCAGACATTG  
7501 TTTAATGAAC TGGGGATATA GTGGTGAACA AACTGACAG CGTTCTTCAT  
7551 TGTATTCTCA AAACCCCTCC TATAGTAAAG AGGTCTGTGT GTGTGTGTAG  
7601 GTGCATGGGG AATAAAAAAT AATAAGCAAA TAATGAACAG GGTAATTTCA  
7651 AAAAGCAGAA AGAGCTATTC AACAAAACCTA CCTGCCTTTT ATTAGATGAA  
7701 ACTCTCAACT CTATGGTTTG TTCTCTCCTG TCAATCTGTG TAAATGCTGT  
7751 CAGCCTGTTT TCCTTATCAC CCTGGCCACG ACTTCTGTCT TTTCTGCTTG  
7801 GTCCTGTAGA CTCTAACCCA AGGCTCATTC TCTGCCTGGC TATCTGCCTT  
7851 CTGTGGCTCT TTGCCACTAC CTACATTTTC TGTGTTGCAC AGGGAAGGAC  
7901 CATTCCCTGT GGACCATAAA ATTCTCTTTT TGAAAGAATT CATTCTTGAT  
7951 TGGGCCACAG CACATCTTGT GAAACAGCAT TAGACATTTG CCACTGCTCA  
8001 GCAGCTCTGG GGGAAAATGT TTAGTGAGAA GCGTACAGTA GTTTTTTTGA  
8051 CTAACCATGG TGCAACCTCC TCCCAGAGGG AAACCTATGA GTATTTCAAG  
8101 GACATGTGAT GGTCTGTTTT TGTCCCCAGT ATCTGACATG ATGGGTAGTG  
8151 TAGACCAAGA GCTTACAGAT AATGGCTAAA TTAAATTTTC TTTTGAATT  
8201 TTAATATTCA ACTTTTTAGG GTACCCAATC TCCATATTTA GGAAAATAAA  
8251 TTACATAAAA AGTGGAGAGT TTTTATTGTG AAACCTGCACC TCCATATTCC  
8301 CAGTGGTGCA GGATGAGGGA GCACAGGTGT TGGTCTGGGG AAGCCAGGGC  
8351 CCTCTGTGGT TCTGGAGGGT GAGGATTAAG AGGAAGCCTT AGATAGTATT  
8401 TATGAGTATC TGCTGACTTC TCTCTGGGAC CCAAGATCAC TGAACCTTTG  
8451 CCTATTTTGA GATCATCTTT CCAATCCAGC CACTAACAGC TGAAGGATAG  
8501 GCTTGCCCTG GAGCCATTGT AGTGGTTGGA TGAAGATAAA AGATAAAAAA  
8551 CTGTGAGGGG AGGTGTCACA GAAGAAAGGG CCCATGTGGG CAGATTTTCA  
8601 TTCAATTCCT AGTCTTTATT ACAGCAATTC TCCAGTGCTG CAACCTTAGA  
8651 AAAGGATTCC TACAACACAA TGTAGGTACC CATCAGCAGC AGATTGGATA  
8701 AAGAAAATGT GGTACATACA CACCATGGAA TACTATGCAG CCATAAAAAA  
8751 GGAGCAAAAT CATGTCCTTT GCAGCAATAT GAATGCAGCT GGAAGCCAAT  
8801 AACTTAAACG AATTATTGTA GAAACAGAAA AACAAATACT GTGTTCTCAT  
8851 TTACAGGGGG AGCTAAACCT TGGGTAAATG GGGCATAAAG ATGGGAACAA  
8901 TAGACACTAG GGACTCCAAA AGGGGGGAGG GAGGGAGGAG GGCAAGGGCT  
8951 GGAAAGCTTC CTACTGGGTA CTTTGTTCAC AACCTGGGTG ATGGCACGAT  
9001 TAGGAGCTCA AACCCTAGTA TCACACAGTA TACCCTTGTA ACAAGCTGAT  
9051 GGTGTAACCC CTGAATCTAC AATAAAATTA TTTTATTTTA AAAAATCATT  
9101 ATAAGGATT TTA AAAAGAA GGATTCCTAG ACAGGTGCAG CCAAACAATT  
9151 TTTTAAAAAT GTTGGCAGGC CGCCACCGCC AGTCACTTAT GCTGCAATAG  
9201 CCCATGTCCC AACATTCCCA ACCTACTTCT CTCCAAAAGA GAAGCTATAC  
9251 TTTTCAGATGG CCCTGTGCTG GGTCTCCCTT GGAAGTTTCT GGGGAAAGGG  
9301 GCTTGAGTTG CCCCAGCTGG ACTCTTCTCT GAGTGGGAGC CGGGGCTTCT  
9351 GATCAGACGT GAGTGAGGCA GGAACCTCCG GGTCTCCAG CGCAGCCCAG  
9401 AGTGCGGTCC CACGCAGGTC CCGGCTCCTG CGCGCTCGCG CTTTGGCGCT

FIGURE 3, page 3 of 42

9451 GAAGCCGTTA GGATGAGCCC TCTCCTTCCA GAGCTTTAAC CGATGAAGGT  
9501 GCATTGTGTT TGGCGCCCTT GAGGAGGATG CTGTCTTAGG CCTCTTCCCA  
9551 CTGGACGTGT GTGGTGGGCA GAGATCCCGT TCGTCGGTCG CACTTCCACC  
9601 CCGCTGGGGC TCACTCAGGC CGCGGAGCTG CGAGGGAGAC ATCCTCGATG  
9651 GACTCCCTCT ACGGAGATCT CTTTGGTAC CTGGACTATA ACAAGGATGG  
9701 GACCTTGGAC ATTTTGGAGC TTCAGGAAGG CCTGGAGGAT GTAGGGGCCA  
9751 TTCAATCTCT AGAGGAAGCG AAGGTGGGTC TCACTGGGGC TGTAATCAGA  
9801 GAGACGTTGG GGCTGGGAGC CCTGGAGAGG CATTGGGCAG AGAGGGCAAA  
9851 ATTTACATGT TGTCAAGCTT GACCTGGGCC CACTGCAGTG TTCAGGTGGT  
9901 TGACCAGCGT TACCGTTTAT TAAGAATAAC AACACAGTA ACACATTTCT  
9951 CAAGTATTTT TCTCCGTTTT CTCCTTGGCT GTAGTAAAAT CTCCAAC TTC  
10001 AGATTGCTCT CAAGATGTTG GCTACATACA GCCTTGTCTT AGGAGTCACC  
10051 TTGTTCAATG TGCTCACCTG TCATTAGTCA CCCAGAGGGG CGTCTAGGCT  
10101 AAAGATGCGC CCTCCCCAGT TCAGAGAACT GGAATAATCA CTCTACGTGT  
10151 ATTTGGGAGT GGGGTGGTGA TTGGAAATTT TCTGATGTTA TGTTTTGGTT  
10201 TCTGTTCCCTG GAAGGGGGCA GTGGAAGTGG CTTTTACTCT CGGGTTTCAC  
10251 TAGTGCTGAG GTTTCCTCAT AATATGCCTT AATTGATAGA CCCTAGTTAT  
10301 CAGTACCGAG CTTAGGCTAA CCCTTCTCTT CCCCAGAAGG CTAACCTACA  
10351 GGCTCCTTCT CAGCATGTTG TGCTTCGTAC ATACTCCTAT TGCAGTATTT  
10401 CCAAGTCATT TTTCAATTGG AATTTATTAT TGTATATAAT AATTACTTTA  
10451 TAAGTATATT TGCTCTTTGG ATGTTTGACC CGGTAGACTG GGAGATCATG  
10501 AGCATGTGGA CTATTGAGTT TATTTTGGAT AATTGGTACT TCGTGCCCAA  
10551 AAAACTGTCA GTTGAGTTCT GTCATGTTGA AATTTAGTAA AACTCTTTCT  
10601 ATTAGCCATG TGAAC TTTGG GAATATTGAA GCATCCATTC AGTCATGGGT  
10651 CAGTTC TAGT TTGAGCACAT TCTATATTCC AAGCCCCATA CCCTGGTATC  
10701 CTCATCTGTT ATATCAGAGG CCTGGACTGT GTACTTTCTG TGGACCAATT  
10751 CAGTCCAAAA TGTTATTTCT GCAAAGCTTA TCTGGATTTT TAATTCCTAG  
10801 AAAAAAGCAG TGTTTCTCCT TTTAAAGTTA AGTGTCTTGT TTCAGGTGCA  
10851 GTGGCTCATG CCTGTAATTC CAGCACTTTG GGAGGCCAAG GCAGGTGGAT  
10901 CACTTGGGGT CAGGAGTTCA AGACCAGCCT GGCCAATATG GTAAAACCCC  
10951 ATCTCTACTA AAAATGCAAA AATTAACCGG GTGTGGTGGT GGGTGTGTGT  
11001 AGTCCCAGGA GGCTGAGGCA GAGAATCAC TTGAGCCTGG GAGGCAGAGG  
11051 TTGCAGCAAG CTGAGATTGC ATCACTGCAC TCCAACCTGG GTGACAGAGT  
11101 GAGACTCCAT CTCAAAAAGA AAAAAAAAAA GTTAAGTGTT CTTCATATTT  
11151 GTTTAAAGAC ACTCTTATAT TTAGATTTGC AAGTGTAAGT TGTATTTGTT  
11201 TATTTGATAC AAAC TAGCCT TTCATAAGAA ATTCTGGGTT AGCTATCAAG  
11251 TCGAATCTTT TGAAACACAT TTCTTCCTTA TTGAAACAAA AGGTTTGTAG  
11301 AGCTGTCTTG CATTTTGGC AAGGACGCTT TGTGTACCTA GTGGTGACTG  
11351 AGGAGGTTT ACATGTCAA ACCCAAGGGA GGGGTGTCCC CAGAGAATTC  
11401 TGCACCAACC ACACAGAAC TTTCTGTTCA GAGGAGCACC ATTGTGACTT  
11451 TTCCTCAAGT GGCAGTCACA TCGTTAGGAG GTTTTGATGT GAGGTCTCTT  
11501 CCCACACGTC TCCACCTCCC CAGTAGGAAA ATTTGTTTAT ATAGACAAAA  
11551 CTCAACTGAT TAAAAAATA AAAAAAGAAAT GATACTTACA TTGTCGTGTT  
11601 AAGATACAAA AGCAATAACT TTTTATTGTG AAAATAGTCT GTTTTGAAC  
11651 AATATATTGT TTTGTTTTTT CCTGTGAAAG TTGAGAAACT AAATATACGA  
11701 AGAGATAATG GTCAGACCAT AAATAAAAAAT AGAACTTTGA CTCAAAATTT  
11751 ACAGCAGTCT GCCCAGAAAA CCAGCCCTTT ATCTAAAAATA AACAGACCAG  
11801 GAAACCAGCC TGTTATGTCA GACTTATAGG AAGTCAGGTT GCTATCTCTA  
11851 GAGACAATAC ACAAAGCTAT GCAATAACTG CTGTAACAGC CCCAAATGGT  
11901 CAGAATTTGA TTAATAACCG ACAGCCCCC TAATTTT TTTT CTTCACTNNN  
11951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNTT  
12001 ACCGCTTGCT AGAAGTGGG CCTTGGGTCA TGTTATTTAA TGCCCTGGAGG  
12051 CCTCAAAATG TAACTAGGTA ATGGTAAAGAC CTACCCAGTA ACTTAGCATA  
12101 AATAGTAAAT TCATTCATTT AATGTTTTCA AACAGTGCCA GACATTGTTT  
12151 AATGAAGTGG GGATATAGTG GTGAACAACA CTGACAGCGT TCTTCATTGT  
12201 ATTCTCAAAA CCCTCCCTAT AGTAAGTAGG TCTGTGTGTG TGTGTAGGTG  
12251 CATGGGGAAT AAAAAATAAT AAGCAAATAA TGAACAATAA AATTATTTTA  
12301 TTTAAAAAAA AAGAAATGAT ACTTACATTG TCGTGTAAAG ATACAAAAGC  
12351 AATAACTTTT TATTGTGAAA ATAGTCTGTT TTTGAACAAAT ATATTGTTTT  
12401 GTTTTTTCTT GTGAAAGTTG AGAAACTAAA TATACGAAGA GATAATGGTC  
12451 AGACCATAAAA TAAAAATAGA ACTTTGACTC AAAATTTACA GCAGTCTGCC  
12501 CAGAAAACCA GCCCTTTATC TAAAATAAAC AGACCAGGAA ACCAGCCTGT  
12551 TATGTCAGAC TTATAGGAAG TCAGGTGCT ATCTCTAGAG ACAATACACA

FIGURE 3, page 4 of 42

12601 AAGCTATGCA ATAAC TGCTG TAACAGCCCC AAATGGTCAG AATTTGATTA  
12651 ATAACCGACA GCCCCCTAA TTTTTTTCTT CACTTCCAAC TTAGGACGAA  
12701 CCAGAGAAAG CTAAATATGC ACCACCTACT AATCAAATAG GGTGCCGCGT  
12751 TTCTAATGAA CCCTCCTACA GCTTCCCCAG GCCAGCAGCC CCCAATCAGG  
12801 AAACGCCCTGA AGCCTTCCCT TTTTCTCACT GTAAAGCTTT CCCACTCCTC  
12851 TGCCTGGCTT TGAGTCTCTG TCAATACACA AGTGAGGGTG TCTGACTCCC  
12901 TTGCTATAGC AAACCTCGGC CAAGTAGATT TTACTTTTCT CATTTGATTG  
12951 GTCTTTTATT TCTAGAAGGA ACATACAAGA AAATTTAAAG GGGAATCCAT  
13001 TCCTAATCTT TCATATTATA GTAGTCCCTT TTTATCTGCA GGGCATATTT  
13051 TCCAAGACCC CCACTGAATA CCTGAAACTG TGGGTAATAT TGAACCCAT  
13101 ATATACTCTC TCTATATATA CATATATATA TATATTTTTT AATTTTTTTT  
13151 TACTTTTATCT TTAATTAGCT TTAGCTCTTT TTTTTTTTTT TGAGATGGAG  
13201 TCTCACTCTG TCACCCAGGC TGAGTGACAG GGTGCAGTCT TGGTTCCTG  
13251 CAACCTCTGT CTACCCGGGT CAAGCAATTT CTTGTGCCTC AACCTCCGGA  
13301 GTAGCTGGGA CTACAGGCGT GTGCCACCAC TTCTTGCTA ATTGTTTAA  
13351 ATTTTAGTAG AAACGGGATT TCACCAAGTT GGCCAGACTG GTCTCGTACT  
13401 TCTGACCTCA AGTGATCCGC CCACCTTGGC CTCCCAAAT GCTGGGATTA  
13451 CAGGCGTGAG CCACCATGCG CCCAGCCATA GACTATATAT TTTTGATCTG  
13501 ATAACCTGTT CAGCTACTAA GTGACTAACA GGCAAGTAGC ATCTATAGTG  
13551 TGGATATGCT GGACAAAAGG ACATTACCTT CCTGGGCAGG ATGGCACAGA  
13601 ATGTTGAGAG ATTTTATCAT GCTACTCAGA ATGGTGTGCA ATTTAAACT  
13651 TATGAGTTGT TTGTTTCTGG AGTTTTCCAT TTAATAGTTC AGACCATGGA  
13701 TTGACCGCAG GTAAC TGAAA CTGTGGAGAG TGAAACTGTG GATAAGGGAG  
13751 GACTATTGTA TTGTTAAGTC AGACTCATTA GGCAATCATA ACTCTTGATT  
13801 TGCCATCAGA AATGCTGCAG AAATATGGGT TAAAAAAAT TGTTCAAAA  
13851 TAGGGTCAGG GATGTCCTTT AACTTGTTAC TTCCAAAATG TTAGTGAAAA  
13901 CTGTGGCCCC AAAGAGTGAA AGGAACAAAT GACTAAGAGA AAATCTGTT  
13951 TTCAGGATGA CAGATTAAAA AAGAAGCAAC TTGCTGAAAC ACTGAAAATC  
14001 TCTCCACTTG TAAGATAACA CAAAAC TGGC TAAAAC TGGT TGGAATGAAT  
14051 ATGGCCAACT CAAGTCTGCA CAGAACTAAC TTGGTGATGT TACAGCCCAA  
14101 ATTTCCACCA CATATTTTAT ACTAACTCCC CCCGGATTTT CACACATGAT  
14151 CTGTGAGGTA GCATGAAGAG GTAAC TATAGC ATGCCTAAG ACTTGGGAGA  
14201 CCTCCCCATT TCCTTCCACC AATCACCAC TAATCCCAGA ATCCGCCCCC  
14251 AAACCTTTTC TAATAACTAC CTTAAAGCCA GCATAGGGAG ACAGATTTGA  
14301 GCTGGACTCC TGTCTTCTTG TGGGTCACCT TGCAATAAAA AGCTTTTCTT  
14351 TTCTCAACAC CTGGTATTAT AGTATTGACT TCTAGTTTAT CGGGCAGCAA  
14401 GCCCCTTTTG GTCGGTGACT ATTCTTGTTT GCTGATATTT CCATTGGCCA  
14451 AAATATAAAC CTCTTAGATG AACTTTCAGT ACGTAAATGG CGCCACAGAA  
14501 TGCTGTGACA TTTTCTCTTT GGATTATAGC AGGTTACTTT ACTGAATACC  
14551 GTAGGCAGTT ATAACACACT AAGTATTTGT GTATCTAAAC ATAGAAAAGA  
14601 TACAGTAAAA ATATGGTAAT TTTTTTCAAC TTTTAGTTGA GATTGGAGG  
14651 GTATGTGAC ATTTGTTACA AGGGTATATT GCATGATGCT GAGGTTGGG  
14701 GTACAATTGA ACCCTGTCAC CCAGGTAGTG AGCATAGTAC CCAATCGATA  
14751 ATTTTTCAAC CCTTGTCAT TCCCTCCCCG TTCTTGTAGT CCCAGTTTC  
14801 TGCTTTTCCC ATCTTTATAT CCGTGTGCAC CCCATGTTT GCTCCCATGT  
14851 GTATGTGAGA ACTTGTGGTG TTTGGTTTTC TATTTCTGCG TTGATTCGCT  
14901 TAGGATAATG GCCTTCAGCT GCATCCATGT TGCTGCAGAG GACGTGATTT  
14951 TATTCTTCTT TATGGCTGTG TAGTATTCCA TGGTGAAAA TATAGTACTA  
15001 TAACCTTACT AAATCACTGT CATATATATG GTCTATCATT GACTGAAATG  
15051 TATACAGTGC ATGATATATA TATATATATA TCTATAATGT CTTATCCATT  
15101 TCGTGATTA TGAGATTGA TTGCTAATAT TTTATACAGG AGTTTGTGAT  
15151 CTTTTTCACT AGTTGACATT GCTTGTAAAT TTCTTTTTTT TGTGATGTCC  
15201 CTGTTAGGTT TTAGAATCAA GTGTATACCC GCCTCATAAA ATGGGTTGGA  
15251 AAATGTTCCC ACCCTTCTG TTCTCTGGAA AATTGGTGTT TTTTCTTAA  
15301 AGTTTGGTAG ACATTATTGT TAAACCATG GGGTCCTCGA TTTTCTTCA  
15351 TGGAAATGTT TTCAAATTAC ACTTTAAATT TCTTTAAAT CTGAGTATAG  
15401 GGCTATCAGA CTTTCTGCTG TCTTATGTCA GTTTTAAATA AGTTGTTTTT  
15451 GTAGGCGTTT GTTATCTCAC TTTCATATTT TTGATATAAA GCTTTTCATA  
15501 ATATCATTA TGTCTATAGT GTCTAGTAGT TTCCATCTTT ACTTTCTGAC  
15551 ATTGGTTATT TGCCAGTTTT AGGAGTTTAT CAATTTTATT AGTCTTTTCA  
15601 AAGAACCATC TTTTGGCTTT GTTAATCCTC CCAATGGTGT GTTTTCTTTC  
15651 TCATTACTTT TTGCTCTTTA TTTCTTCAA CTTCTTTTTT GCTTAATTTT  
15701 AAAATAATTT CTTGAGATTG AGATAAGCCT CAATGATGGG TCACCGATTT

FIGURE 3, page 5 of 42

15751 CCAGTCTTTC TTCTTTTCTA ATTATGCATT TTAAACCAGA AATCTTTCTC  
15801 TAAGTGTAGC TTTAGTTGCA GCTCACAAGT TTCAGATCTG TCTCTCAGTC  
15851 TGGAGGTTGG AGATCTGACC ATGACCATGA AACCATCCAG TCACAATGTG  
15901 GCATTATTTT TTTAATTTTT TTTTTTTTTT TTGAGATAGA GTTTCACCTC  
15951 TATTGCTTAG GCTGGTGTGC AATGGTGCGA TCTCGGCTCA CAGCAACCTC  
16001 CACCTCCCAG GTTCAAGCGA TTCTTTTGCC TCAGCCTCCC AAGTAGCTGG  
16051 GATTACAGGC ATGCGCCACC ATGCCCAACT AATTTTGTAT TTTTAGTAGA  
16101 GATGGGGGTT CTCCATGTTG GTCAGGTTGG TCTTGAAGTC CCGACCTCAG  
16151 GTGATCCGCC CACCTCAGCC TCCCAAAGTG CTGGGATTAT AGGAATGAGC  
16201 CACTGTGCCC GGCCCAACTT GGCATTATTT ACCCAGAAGA GCATGACCAT  
16251 GAGAACAGTA GAATTTGTAA GCTTTGAGTG GGTGACTATG AGTGTCATAA  
16301 TAGGTGAGTA GGTATATTTT TGGGTGGTGG TAGGAGAGGG CTTACAGTTT  
16351 GCTATGACAG CTTTTTATAT GGATCATCCT TAGTAAAAGA TTATTTAATT  
16401 TTTGAAATCA AAGGGGAAAA CACTAGTTTA GGCTTTCTTC TTTCTTTCTT  
16451 TTTTAGAGAC AGGGTCTTGC TCTGTCACCA GGTTAGAATG CAGTGGTGCA  
16501 ATATTGCTCA CTGTAACCTC AAATTCCTGG GCTCAAGTGA TCCTCCTACC  
16551 TCAGCCTCCA AGTAGCTAGT ATTTACAGGC ATGCACCAAC ACATCTGGCT  
16601 AATTTTAAAA ATTTTATATG GAGATGAGGT CTCACTATGT TGTCCAGTCT  
16651 GGTCTTGAAT CCTGACCTCA AGTGATCCTC CCCCATCAGC CTCCCAAAGT  
16701 GCTGCAATAT TTTAAATCCT GTGGTAGGTC AAGTGGTTGT CTTCTATCTT  
16751 GGGGTTTATA AAGTACATGT CAAGAAATTT AGGGTATGGT TAGATTAGCT  
16801 TTAAAAATGT CATGTTTAT AAAAAATCAAT GCATCATTTT TCTGATTGAA  
16851 AATTTAACAC AAGACTCAGA ATCTTTTTCG AGTAGTGGAA TTACTTTTAT  
16901 TATAGATCTT TGCGATAATG AATGATGATA CATCTGGCCA AAAATAGGTA  
16951 CTATAGTCTT TTAGGAAAAC AGCTAATCTG CTTGAAATAT GTGTAGAAAT  
17001 AATTTAGTGC ATCAGCCCAT ATTGGCAATA ACTTCTCTCT AATTTTTTTT  
17051 TATAGAAAAT TTTTACTACT GGAGATGTCA ACAAAGATGG GAAGCTGGAT  
17101 TTTGAAGAAT TTATGAAGTA CCTTAAAGAC CATGAGAAGA AAATGAAATT  
17151 GGCATTTAAG AGTTTAGACA AAAATAATGA TGGTGTGTCT TTCTTTTGTA  
17201 TTTATCACCA GCTATGAAGA AGCATTATATC ATGCTTTCAA GAGTCTAAAA  
17251 GGATGCTTAT TTAATCTCTC TGGTTTTAGA TGATAATTAT TATTTGTGTT  
17301 AATACCTTTT TATTAGTAAAG TGAATTTTAT GTAGAGTTTA TATTATTTAG  
17351 TGAAGAAAAC TTATAGATAG CTTTTCTTTT TCATTACTTT GAAATGTAAT  
17401 GAATTACATT TCTGAATTAA AACTGTGGG CAGGGCCTGT TGTAAATGTT  
17451 AACTATGGAA CATTATGCTG ATTTGAGTTA AACCTGTAGG TTAAAAATAA  
17501 TAATTATATT TTCTTGTCCT CTGGGTAAAA TGAGATTTCT TTTTATTTGT  
17551 ATAGAAGAAT GACAGTTGTG TCATCTAAAA TTTAAAAAAC TTTCAGATTA  
17601 TCTTGCACTC GTTAGTTTTT TTGGAAGAAT TAATTTAGAG AAGATATCTC  
17651 TGATCCTGGA AATTAGGGAA AAATAGCATA TAAACGTTTA AGTGTGTACC  
17701 TTCTGGTTAA GATTATGACT TCTATATTTT GATTAAATAGG TTGGAGTTTG  
17751 TCTTAATCTG TTTTCTGTTG CTGTAATGGA GTACCACAGA CTGGGTAATT  
17801 TATGAAGAAA TGAAATTTAT TTCTTATAGT TCTGGAGGCT GGGAAAGTTCA  
17851 AAGTTGAGCC GAATCTGGTG AGGGCCTCTT ACTATGTCAT AACATGCTAG  
17901 CAGGCATCAC AGAGCAAAAT CACTACCTCA GATCTCTCTT CCTCTTCTTA  
17951 AAAAGCCACT AGTCCCATCA TGGGGGCCCT ACTCTGAAGA CCTATCTAA  
18001 TTCTAATTGG AAATAGGGTC TTGAAGCCCT CATCACTAGA GGTAACCTTT  
18051 AACAGGAAGA GAGAATTTAT AAAAATTATA ATGCAGCACC AAATCCCTCC  
18101 CTACTTGTGA ATAGTCAAGG TCATTTTATT TACAGACTTG TTATTAAAGA  
18151 AACAGGTAA ACAAATAGAT TGAGAGGAAA TGTGGTTCAT GTCTGAGATC  
18201 AGCAAACCTT TTTGTCCAGA AGTCCAGATA ATAAATATTT TAGCTTTGTG  
18251 GGTCAATGTT TCTCAGTTGT AGCTACTTGT CTCTGCTGCT GTACCTCAAA  
18301 AGCAGCCATG GATAATATGT AAATGAATGG GGATGACTGA TTTCCAATAA  
18351 AAACCTTATT TACAAAGATA GTTAATACAC CTTATTTGGC TTGAGGGTTA  
18401 TAGTTTGCCA TCCCCTGATT TACAATGAAT ATTAAGTTT AATTCAAAGC  
18451 AAGTTCTCTC AAACAAACAA ACTAAACTCT AGATGATTTT GAAGATTATT  
18501 CACATCTGTG ACTCTCAGCC AGGAAGAGCT GAGTTTGGGT TGGAAGTAG  
18551 TACTATTGGA ACATTTGTTG CCCATAAGCC TTACAATATA TGCCCTAAG  
18601 TCTAGCCTTA GTCCAGTCTT CTAGCAAAAC TCAGTTTCTT TTCTTCTCTG  
18651 CAACTTTTCA TTCCAACATC GACCTCTGCT AGTTTCAGATT GTCTTGCAGG  
18701 TCAGATTGTC TGTGTGCTGC TATGGTAGGC AGTAGCTGAG AGATGGAGCT  
18751 ACCTTAAGAT CAATTGCCAG ATAATCAGAG GTCAATTATC CCAGTGCATA  
18801 AGTAGTGAC ATATCAATTG TTCATTTTAT AAAATTCTAA ATGAACCAGA  
18851 GGCAATAATT AAAGATGAAA TTTTGATGGT ATATTTGTAG GAAATCTACA

FIGURE 3, page 6 of 42

18901 CAATGTTTCC CTAATTTCCC ATGTTTGTGT ATTTTAAAAC AATGTGGCAT  
18951 TATTGGTTCA TATTTTATT TTTTAGACTT CCTTAATGCA AAACATATAC  
19001 AGTTGATCCT CATTATTTGG GGATTCTGTA TTTGCAAATT TGCCTACTCA  
19051 ATAAAATTTA TCCCCAAAGT AACCCCAAAA TATATACTCA CAGTACTTTC  
19101 CCAGGCATT CAGACATGC ACAGAGCAGT GAAAACTTG AGTTGCTCAG  
19151 CATGTACATT CCTAGCTAGT AGAATAAGGC AATACTCTGC CTTCTTGTTT  
19201 CAGCTCTCAT ACTATTAAC AGCAAGTATC CCTTTCAGG TCTATTTTGT  
19251 GCCAGTTTTT GCATTTTGTG ATTTTGTGTG GTAATTCCT TTTTAAATG  
19301 TTCCCCAAG GTAGTGCTGA AGTGCTGTCT AGTGTTCTTA AGTGCAAGAA  
19351 AGCCATAGCA TGCCTTATGG AGAAAATATA TGCGTTGGAT AAGCTTTGCC  
19401 CCAAATTCAT TGTTAGTGAA TCAACAGCAC ACATTAAATG AAGTGCCTTC  
19451 AAACAGAAAC AGACATAAGA CATGGTTATG TATTAATCAG TTGATGAAAG  
19501 TGTTGTAATC AGAGGCTCAC AGGAACCTAA CCCTGTTTTT CCTGTAGGAA  
19551 CAATGGTTTG GTATTTGCTA ATTCAGTGTT TGCAATGAAT ATAGAAGTTT  
19601 ATGGAAGATG ATTGCTGTGA ATAATGAGAA TTAACCATAT CTCTTAAAGA  
19651 GTGCATTTCT AAAGGAGAAT ATTCAGAAGG GTATTTGCAT AATTTCTTTA  
19701 CTAACAGATG CTGCCCTCTCA CTGCTCTTAC ATGGTCCAGA TTCTCATGCT  
19751 GCTCCTTCCC TCTCCCCAGG AGGATTCTCT CAGAATCCTG TCATCTCCTC  
19801 CAGGGTCCCT TCTCCAAGAA AGTCTATCCT TTCACACTA ACAGTAATTT  
19851 TGGTCTTCCT CTTTTTCTGG AGAAGTCAGC TGTTTATGCT GCTTCAGCAC  
19901 CAGACCTCT CTTACTTTGT TTTGTTTCAT TCTTTTTCAT GTACAGTAGT  
19951 CTTAGGATTC TCATGAGCCT GTGAGCTGCT AGAAGGAAAT ACAGCAGTGC  
20001 TTACATTTAT TGCTTCTATT TTATTTTCTA TTTTCTCTTC CTGCTCTCTG  
20051 ATTGTTCTCC TTCTGTCCAC AAACATGCTC TAATTTCCCT AGTATTAAAA  
20101 ATTTTCTGTC TTTTGTGTG CTTTATCTCT TGCTCCCTTA TTTTACTGCT  
20151 CAGATTTTAA TTTTATTTA TTTATTTTGG AGATGGAGTC TCACTCTGTC  
20201 ACCCAGGCTG GGGTGCAGTG GCGCGATCTC AGCTCACTGC AACCTCCGCC  
20251 TCCCAGCTTC AAGCAATTTT CCTCTTTTAG CCTCCCAAGT AGCTGGGATT  
20301 ATGGGCACCT GCCACCATGC CTGGCTGATT TTTCTATTTT TAGTAGAGAC  
20351 GGGGTTTCAC CATGTTGGCC AACTGCTCT CTAAGTCTG ACCTCAGGTG  
20401 AACCACCCGC CTCAGCCTCC AAAAGTGCTG GGATTGCAGG TGTGAGTCAC  
20451 TGTGCCTGGC CTTTACTGGC CAGATTTTAA AAAGAATAGT CTGTGCTTTA  
20501 GCTCTATTTT CTCACTTACT ACTTCTCTTT AACTCAGTCA TATATGATGT  
20551 TTTGCATAGT AAATGTCTAG TAATTTATTA AAAATGTAGA AATAGGTACT  
20601 TTTAAAATGA ATAGATCCTA CTTTAAATGA ATTTATCTTG GAGTTAGAAT  
20651 ATCTTGATTT GGATTTTAGT TCTGCTACTT CTTAATTACA TTACTTGGTA  
20701 AGGCCACTTG TGAAGTCAGT CTCTTTGGAG GAATATTATT TATCTATAAG  
20751 GCTGTTACAA TTAGTGAATT TTAATAAATG TGTATTTATT TTTTAAATGA  
20801 TTTGTTACAT TTTTAGTATT GATGTTGGGA TAGGCATTTA AGCAAGTCTA  
20851 TAACTCACCT ACATGCATAA TTTTGCCTTA ATCAGTTTAA AGCTTTCTCT  
20901 TAAATGAGAG ATTTGAAATT CATAATTTCT GTGGTTCTTA TCAGTTCTGA  
20951 GTTTTATTTT TTGCCCTTTT TATTTTTTTA AAGGAAAAAT TGAGGCTTCA  
21001 GAAATGTGCC AGTCTCTCCA GACACTGGGT CTGACTATTT CTGAACAACA  
21051 AGCAGAGTTG ATCTTCTCAA GGTAAGCTCT TCATGTTGGT CAACAATTGA  
21101 CTTTCACTTT AATATCCTGC ATTAGAAGTCT TGTGTTTGTA AGTGTGGCTT  
21151 TAAAACACCT CCCTAGTCTT CATTATGTAT ATCCAAGATC TTTTGTCTT  
21201 TTTTCTCCCC ATTCATTTTG TATGTGTACA TTTATCTAAA GTGTAAGAAT  
21251 GGGAAGTGTA AGCTCAGACT GGACTCTTTC TTTCAAGGCC TCAAAGGATA  
21301 GTGGAATGGC AGGAAGTAAG GTTTTAACTC CATAGATGAG GAGCTGAAGA  
21351 GTTTTGGTGT TGCTTTTCTT CCATTTGATT TCTAATGTGA CAGTAAACT  
21401 CATTGATTCA AACTAAGAAG ACTAGCAGAT TCATCAGATT ATTTAACCTA  
21451 GATGTGACTG GAAAAAGGG AAATTACTAA GCTCTCCAAG CTAACAAAGA  
21501 AATACCTGTT TAACTTTTCA GAAAACAGAA ATGCAAATTT GAACCTTATT  
21551 GTCTGGGGCA ATCAGTTTGA CTATTTAAGT CAGACTTTTA TACTCTTAAT  
21601 GTTTTGTTC ATGGGATAGA GCAGTAATCT CTGCAGCCCA GGTGCTCTCA  
21651 AATACTCTGT TGCTATAAAC ACAGGGCAGG AACTGATTTT TTATGATAAC  
21701 GTAAAACAGA AAAGGACAAT TATATTGTAT TAATATTGTT GTGAATATTT  
21751 TCAGTCTCTA CATTGTCTAA AAATCTTCT AAATGGCTTT GTTATTGAAT  
21801 TTATCTCATT TTATATCTGT GCCAACAGCA TTTTCTCCTT TTCTCTCAT  
21851 AATTTCTTTT ACAAACAGCT GCTCAAGAGG AAGGCTCAAA GTCTCAAGGC  
21901 TGAGCACGTA ATGACTTTTG TTAGTACTAG ATGAGAAGGG CTTTCTGAG  
21951 GAAATGAAAA CCTAAACAT GAAAAGAAGA TAAACAGAAT TTGGACAGTG  
22001 AGATATAGAG CATATAATAT TCTGCTCTA AAGTAATATT CTTCTAGGAA

FIGURE 3, page 7 of 42

22051 AGTGAGGGCG TTTCCCTGGC TGTTAGGCCA GAAATCATAT TCCTATATTT  
22101 TCTTTGATAG CTTTAGGAAT AATGCAAATT CTAAGCCCAA GCTTCAGAAT  
22151 AGACTAAGAA GTATTAGCTT AGCTGCCATG ACAAATACC ATAGGCTGGA  
22201 TGCATTAAAC AATGGAATTT TAGTTTTTCA CAGGTCTGGG AGCTGGGAAG  
22251 TTTAAGATGA GAGTGCCAGC ATGGTTGGGT TGTAGTGAGG GCTCTCTTTC  
22301 TGGCTTGCCAG ATAGACCCCT TCTCACTGTA TTGTCATATG GCAGAGAGAG  
22351 AGAGAGAGAG AGAGAGAGAG AGAGAGAGGG GATCTTTCTC TTGCTTTCTA  
22401 TTATAAGGCC ATAGTCCTGT TGGATCAGGG TTCCATTCTT ATGACTTTAT  
22451 TTGACTTTAC CCCCTAAGA TGCTATCTCC AGATATAATC ACACGGTGGG  
22501 TTAGGGCCTC AACATTTGGA TTTGGGAGGG ACACAGCTCA GTCCATAGCA  
22551 AAGGATAATG CAGAGGGTTG GATATTTAAA AGTAGCTACA CAATTTTAA  
22601 TATAAATATT TTTCTAACT TTTTTTTTTT TTTTGAGATG GAGTCTAGCT  
22651 CTGCTTGCCCA GGCTGGAGCG CAATGGTGCG ATCTCAGCTC ACTGCAACCT  
22701 CCGCCTCCCA GGTTCAAGCA ATTCTCCTGC CTCAGCCTCC TGAGTAGTTG  
22751 GGACTATAGG CACGCGCCAC CACGCCTGGC TATTTTTTTT TTATTTTTAC  
22801 TAGAGACGGG TTTGCACCAT ATTGGTCAGG CTTGTCTCGA ACTCCTGACA  
22851 TCAGGTGATC CACCCATCTT GGCCTCCCAA AGTGTGCGA TTACAGAAGT  
22901 GAGCCACCGC GCCTAGCCAG CAGCTTTACT GAGATGTAAT TCACATGCCA  
22951 TAAATTCACT TTTCTAAAGT ATACAATTCA GTGACTTAAA ACATTTATTT  
23001 ATTTTTTAAAT TGACAGAATT ACATGTATTT ATCATGTACA ACATGATGTT  
23051 TTGAAGTATA TGTACATTGT GGAGTGACTA AGTCTAGCTA ATTAACATGA  
23101 TACATCTCAT ACTTAATGAT TTCTGTGGTG AGAACACTTT ACATCCATTC  
23151 TCTTAGTATT TTTCAAGAAT ATAATATATT ATTATTAATT GTAGTCTTCA  
23201 TGTTGTATAG TGGAGCTCTT GAACTTATTC CTCATGTCAA GCTGAAATTG  
23251 TGTGTCCTTT AACACAAACC ATACCCGACT CCCAAAGTAT TCTGCTCTCT  
23301 GCTTCTATGA GATTAACCTT TTCTGATTCC ACATGAGTGA GATCATGCAG  
23351 TATTTATTTG TCTTTACCTG GCTTATTTCA TTCATATTGT TACAGATAAC  
23401 AGGATTTCTT TCTTTTTTTA ATGGCCGAAT AGTTTTCTAT TGTATATGTA  
23451 TAGCACATTT TCTCTCTTCA TGCATTGGTG GACACTTAGG TTGATTCCGT  
23501 ATCTTGGCTA TCGTGAATAG TGCTATAATG AACATGGGAA TGCACATGGC  
23551 TCTTTGACAT ATTGATTTCA TTTTATATAT GTGTATATAT ATATGTATAC  
23601 ACACACATAC ATACAGTGGT GCGATTGCAG GATCATATGG TAGTCTATA  
23651 TTTAATTTTT AAAGGAAGTC CATACTGCTT TCCATAATGG CTGTATTAGT  
23701 TTAACCTCTC ACCAACAGGG TGCAAAAGTT CCCTTTTCTC TACATACTTG  
23751 CCAACACTTG TTATCTTTTG TCTCTTTGGT AATAGTCATT CTAAGTGTAG  
23801 TATGAGGTGA TATCTCATTT TGGCTTTTAT TTGCATTTCT GTGGTAATTA  
23851 GTGATATCGA GCTTTTTTTT TTTTTTGAC TTTGGCCATT TGTATGTCTT  
23901 TGAAAAATGT CTATTGGGGT TTTTTGGTTG TTTATTTGAG GTTTTNNNNN  
23951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN  
24001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN  
24051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN  
24101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN  
24151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN  
24201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN  
24251 NNNNNNNCCG GGGTTCCCGT CATTCTCCCT GCCTCAGCCT CCCCAGAGTA  
24301 GCTGGGACTA CCAGGGCACC CGCCACCAC GGCCCGGGCT AATTTTTTGT  
24351 ATGTTGAGTA GAGACGGGGT TTTACTGTGT TAGCCAGGAT GGTCTTGATC  
24401 TCCTGGCCTC GTGATCTGCC CGCCTCGGCC TCCCAGAGTG CTAGGATTAC  
24451 AGGCGTGAGC CACCGCGCCT GGCCTGATTT CTAGTTTTTT ATTATGTGG  
24501 TCGGAAAAGA AACTTGATAT GATTTTCATC TGCTTAAATT TGTTAAGACT  
24551 TGTTTTGTGG CCTAACATAT GATATCCCCT GGTGCATGTT CCATGTGCAG  
24601 TTGAGAAGAA TGTGTATTCT CTTGCCATTA GGTGAAATGT TTTATGTCTG  
24651 ATCTGTCCAT TTGTTCTAGA GTATAGTTTA AGTCTGATGT TTCTTACTGA  
24701 TTTTCTGTTG AGATGATTTG TCTATTGCTG AAGGTAGGGT GTTGAAGTCC  
24751 CCTACTATTG CTGTATTGCA GTCTCTCTCT CTTTTCAGAC GTATTAATGG  
24801 TTTTTATTTT ATTTTATTTG TTGTTGTTGT TGTGTTGTT GTTGTTTTTG  
24851 AGACGGAGTC TCACTCTGTC ACCAGGCTGG AGTGCAGTGG CAGGGTCTCG  
24901 GCTCACTGCA GCGCCGCTCT CACGGTTCAA GCGATTCTCC TGCCCTAGCC  
24951 TCCCAGAGTC CTGGGACTAC AGGCGCATGC CACCACGCC AGCTAATTTT  
25001 TGTATTTTTA GTAAAGACGG GGTTTCACCA TGTGCGCCAG GATGGTCTTG  
25051 ATCTCTTGAC TTCATGATCC ACCCGCCTTG GCCTCCCAA GTGCTGGGAT  
25101 TACAGGTGTG AGCCACCACC CCTGGCCAAT GTTTGGTATT TATCTTTAGG  
25151 TGCTCTGATG TTGGGTTTCAT ATATATTTAT AAAAAACAAT AGCTACATAA

FIGURE 3, page 8 of 42

25201 CTTATTAAGG GATATGCAAT ATAAAATATA TAAATTGTGA CACTGAAAAT  
25251 TTAAAATGGG AGGAGTGGAG TAAAAGTACC TTCATATAAC TTACTATTAT  
25301 ATCCTCTTAT TGAATTGACC CTTTTATCAT TATATAGGAA CTTTGTCT  
25351 CCTTTACAAC TTCTGACTTA AAGTTTGTTC TATATGATAT AAGTAAAGTT  
25401 ACTCCTGCTC TCCTTTGGTT TCTGTTTCCA TGGAATATCT TTTTCCATTC  
25451 CTTCAACCATC AGTCTGTGTG TATTTTACA GATGAAATGA GTCTGTCATG  
25501 GGCAGCATAT AGTTGGATCT AGTTTTTTTA ATCCACTCAG ACACTGTGTT  
25551 TTTTGATTGG ATAATTTAAT CCATTCATGT TCAAGGTAAT TATTGATAAG  
25601 TAAGGACTTT GTACTACCAT TTTGCTTATT GTTTCATGGT TCTTTTATAG  
25651 ATCCTTTTAT CTTTTCTTCC TCTCTTGCTG TCTTTTTTTT GTGGTTAAGT  
25701 GATTTTCTCT AGTGGTATGT TTTGATTCTT TGCTTTTTAT TTTTGTGTA  
25751 TCTCCTATTG GTTTTGGTT TGTGGTTACC AAGAGGTTAC AAAAAACATC  
25801 TTAAGAGTTA TAATAGTTTA TTTTAACTTG ATAACCTAAT TTTTATGCA  
25851 AAAACCCCC AAAACAAAA AATCTACACT TTTACTTAAT CCCCTGAAAT  
25901 TTTGAATTTT TGATGTCACA GTTTACCTCT TTTTCATATTG TGTATCCCTT  
25951 AAATTATTGT AGCTATTATT ACTTTTAATA GTTTTCTCTT TCCTACTACA  
26001 GATGTAAGTG ATTTGCATAC CATCATTAACA GTATTATTTT GAATTTACCT  
26051 GTGTACTTTT TTTTATCAGC CAGTTTTATA CTTTCAGATG TTTTGTGTT  
26101 ACTCATTAGC ATCTTTTCTT TTCAGCTTGA GGAGCTCCTT TTACGTTCT  
26151 TATAAAATAG GTGCGGTCAT GATTATCTCC CTCAGCTATT GTTGTCTGG  
26201 GAAAGTATCT CTCCTTCATT TCTGAAGGAC ACTTTGCTGG GTACATTACC  
26251 CTTGGTTGGT ATTTTCTCC TTGAACGCTT TAAATATATC ATCCCTTCT  
26301 CTCCTGACCT GTTAGGTCCT TGCTGACCAG TCTGTTTCCA ACCATATTGG  
26351 GACTGCTTAT TATGTTATTT GCTTCTTATC TTTTGTGTTT TTCAGGATCC  
26401 TCTCATTTG TTTGATTTT GATAGTTTGA TTGTAATATG TCTTGGGGTA  
26451 GTCTTGTTTG GATTGAATCT GATTAGAGAC CTTGGACTTT TCCTGCATGT  
26501 AGATATTTAC CTCTTCTCC AGGTTTGGAA AATTTTCTGT TACTGTTTCT  
26551 TTAATTAAGC TTTTACCCC TTTTATCTTC CTTTCTCCTT TCTTCAACTC  
26601 CTGTGACTCA AAACCTTGCT CTTTGTATGC TGTTCCATAA ATCTTGTAAAG  
26651 CTTTCTTCAT TCATTTTCAT TCTTTTCTT CCTCTGTGTA TTTTCAAATA  
26701 ACCTGCTTTT GAGTTCATAG TTTCTTCTT CTTCTTGATC ACTTCTGCAG  
26751 TTGATGCTCT CATATTGCAT TTTAATTTT TTCATTGTAT TTTTCAAGCC  
26801 CATGATTTCT GTTTGATTTT TTCTTTTATT ATTTTCATCTC TTTATTACCT  
26851 TTCTCTTTGT GGTCACTCGT TATTTTCTTA ATTTTCATTGA ATTGTTTCTT  
26901 TGTATTTTCT TGAAGTTTGC TGAGCTTCTT TTGAATTCTA TGTCAGTTCA  
26951 TACATCTCTG TTTCTTTAGG GATGGTCGCT GGTACTTTAT TTTGTTTCTT  
27001 TAGTGGTGTC ATTTGTTCCT GATTGTTGTT GATGTTTGTG GCCTTGTGTT  
27051 TACATCTGTG CATTTGAAGA AGTAGGCACT TATTTTCAGT TTTGCAGACT  
27101 GGCTTTGTCT GAGAATGCCC TTCAACAGTC AGCCTGTCTA GAGATTCTTT  
27151 AATATTTAAT TAAATATCTT TAATATTTTG AAGAAGTTCC AAATTGTTTC  
27201 TAAAGTGGCT GCACCATTTT ATAATCCCAG CAGCAATGAA TGAAGGTTTC  
27251 AGTTTCTCCA TAGCTATATG AATACTCATT ACTGTCTGTC TTTTCATTTT  
27301 TTGATTTTTA TTTTTTTTTT GAGAAAGGGT CTTGCTCTGT CATCCCATCT  
27351 GGAGTGCAAT GGCACAATCA TGGCTCATTG CAGCCTCAAC TTCCCTGGCT  
27401 CAATTGATCC TCTCACCTCC TGAGTACCTG GGACTACAGG CATTGTACCA  
27451 CAATGCCTGG CTAATTTTTA TATTTTGTG AGAGATGTGG TTTTGCCATG  
27501 TTGCCTGGTG TATTAGTCCA TTCTCATGCT GCTATAAAGA ACTGCCTGAG  
27551 ACTGGGTAAT TTATAAAGGA AAGAGGTTTA ATTGACTCAC TTTTGTCTGG  
27601 CTGAGGAGCC CTCAGGAAAC TTACAATCAT GGTGGAAGGG GAAGCAAACA  
27651 CGTCCTTCTT CACATGATGG CAGGAAGAGC AGTGCCTAGC AAAGAGGGAA  
27701 AAAACCCCTT ATAAAATAAT CAGATCTCAT GAGAAGTTAC TCACTATCAT  
27751 GAGAACATCA GAATGAGGT AGCCTCCTCC ATGATTCAAT TACCTCCAC  
27801 TGGGTCCCCT ACCTGACATG TGGGGATTAT TGGAACATA ATTCAAATG  
27851 AGATTTGGGT GAGGACACAG CCAAACCATA TCATTTTTCG CCTGGTCCCT  
27901 CCCAAATCCC ATGTTCTCAC ATTGCAAAAC ACAATAATGC CTTTCCAGCA  
27951 GTCCCCCAGC GTCTTAATCT ATTCCAGCGT TAACCTAAAA GTCCAAGGTT  
28001 TCATCAGAGA CAAGGCAAGT CCCTTCTGCC TATAAGCCTG TAAAATCAAA  
28051 AGCAAGGTAG TTATTATACT TCCTAGATAC AATGAGGGTA CAGGCATTGA  
28101 TTAAATATAC GTTTTCCAAA TGGGAGAAAT TGGCCAAAAT GAAGGGGCTA  
28151 CAGGCCCAA GTAAGTCCGA AATCTAGTGG AATAGTCAAA TCTTAAAGCT  
28201 CCAAAATGAT CTCCTTTGAC TCCACATCAC ACATCCAGCT CATGCTAATG  
28251 CAAGAAGTGG GCTCCCATGG CCTTGGGCAT CTGCACTCCT GTGGCTTTTC  
28301 AGGGTACAGA CCCCTTCTG GCTCTTTTCA CAGGCTGGCG TTGAGTGTCT

FIGURE 3, page 9 of 42

28351 GTGGCTTTTC CAGGTGCATG GTGCAAGCTG TCGGTGGATC TACTATTCTG  
28401 GGTACTGGAG GATGGTGGCC CTCTTTTCAC AGCTCCACTA GGCAGTGCTC  
28451 CAGTGGGGAC TCTGTGTGAA GGCTCCAACC CCACATTTCC CTTCTGCACT  
28501 GCCCTAGCGG AGGTTCCTCT CAAGGGCTCC ACCCCTGCAG CAAACTTCTG  
28551 TCTGGACATC CAGGCATTTT CATACATCCT CTGAAATCTA GGCAGAGGAT  
28601 CTCAAACCTT AATTCTTATC TTCTGTGTAC CCGCAGACTC AACACCTTGT  
28651 GGAAGCTGCC AGGGCTTGGG GCTTGACCTT TCTGAAGCCA TGGCCTGAGC  
28701 TGTACCTTGG CTCCTTTTAG CCATGGCTGG GATGCAGGGC ACCAAGTCCT  
28751 GAGACTGCAC AAAGCAGCAA GGCCCTGGGC CTGGCCAGG AAACCATTTT  
28801 TTCCTCCTGG GCCTCTGGGC CTATGATGGG AGGGCCCTTC CTGAAGACCT  
28851 CTGAAGTGCC CTGGAGGCAT TTTCCCATTT GTCTTAGTGA TTAACATTTT  
28901 GCTAATTCAG TCTTATGCAG ATTTCTGCAG CTGGCTTGAA TTTTTCCTC  
28951 AGAAAATAGA TTTTCTTTT CTGTACATC ATCAGGGTGC AAATTTGACA  
29001 AACTTTTGTC CTCTGCTTCC TGTGGAATGC TTTGCCACTT AGAAATTTCT  
29051 TCTGCCTGAT ACCCAAATC ATCTCTCTTA GGTTCAAAGT TCCACAGATC  
29101 TCTAGGCGAG GGGCAAAAAG CCACCACTCT CTTTGTCTATA GCATAACAAG  
29151 AGTCATCTTT GCTCCAGTTC CCAACAAGTT CCTCATCTCC ATCTGAGATC  
29201 ATCTCAGCCT GGACTTCATT GCCCATATTA CTGTCAGCAT TTTGGTCAAA  
29251 GCAATTCAC AAGTCTCTGG GAACCTACAA ACTTTCCAC CTCTTTTGT  
29301 CTTCTGAGCT CTCCAAATTT TTAAGAAGTT CCAAACCTTC CCAGTCTTCT  
29351 TCTGAACCTT CTAACCTGTT CCAACCTCTG CCTGTTACCC AGTTCCAAAG  
29401 TCAGTTCCAT ATTTTGGGT ATCCTTATAG TAGCACCCAA CTCCTAGTAC  
29451 CAATTTACTG TATTAGTTCA TTCTCACGCT GCTATAAAGA ACCACCTGAG  
29501 AATGGGTATT TTATAAAGGA AAGAGGTTTA ATTGACTCAC AGTTTCGCGT  
29551 GGCTGGGGAG GCCTCAGATA ACTTACAGCC ATAGCAGAAA GGAAGCAAA  
29601 CATGTCTTTC ACATGGTGGC AGGAAGAAGA AGTGCTGAGC AAAGAGGGAA  
29651 AAGCCCTATA AAACCATCAT ATCTCGTGAG AACTCACTCA CTATCATGAG  
29701 AACAGCAGCA TGGGGTTGAC CACCCCCCAT AATTCAATTA CCTCCCACCA  
29751 GCTGTCTCCC GTGACACATG GAAATTATGG GAAC'TACAAC TCAAGATGAG  
29801 ATTTGGGTGG GGACACAGCC AAACCATATC ATCTAGGCTG GTATCGAAAT  
29851 CCTGGGCTCA AGCAATCCAC CCACCTGGC CTACCAAAGT GCTGGGATTA  
29901 CAGGCATGAG CCACCATATC TGAACCTGCT TTTGATTCT TTTGATTTA  
29951 ACCATCCATT GTTTCTGCTT CTCTAGATAA CCCTGACTAA TATAAATTG  
30001 GTATGAAGTG ATATCTCATG GCTTTGATTT ATATTTCTTT CATGGCTAGT  
30051 GACTTTTTTT GTACTTTTGG GATATTGTTA TTATTATTAT TATTATTACT  
30101 AGTGTTTATA CTTCTTCAGT AAAAGTGTTA GAAACAATTT TTAAGGCAG  
30151 AATGTGACCA GAGTTTCCTG TAGTTATATA ACCATCATGG ACCTTCCCTC  
30201 AAGTGCTAAG CCATTAGTGT TACTCATGTC ACTCCAATG TCAGCTTGTT  
30251 TTCTTCCATT TCACTGTCTC TTTGTGTCCC AAAC'TGAAT TCATGGGAAA  
30301 AACATCTGAA TGGTGCTTAA TATGGTTTGG ATATTTGTCC CCTCCAAATC  
30351 TCATGTTGAA ATATGACCTC CAGTGTTGGA AGTAGGGACT ACTGGGTCA  
30401 CGAGAGTGGA TCCTTCATTA ATGGCTTGGT AATAAGTGAA CTCTATTAGT  
30451 TCATGAAAGC TGGTTGTTGA TAAGAGCCTG GCATCTCATT TCTCTGTCC  
30501 TTCTCTCACC ATCTGACACA CTTGCTCACC TTTTCTTCT AGCCATGAGT  
30551 AAAAGCTTCC TGAGGTCTCA CCAGAACTG AGCAGATGTT GGTGCCATGC  
30601 TTGTACAGTC TGTAAGACTG TGAGCCAAAT AAGCCTCTTT TCTTTATAAA  
30651 TTACCGAGTC TCAGGTGTTT GTTTAAAACA ACACAAAACA GACTAACACA  
30701 GTGTTGATTG AAACAGCTGT GACTGGGTCA TCAGGGTGTA AGAGAGGAGT  
30751 CACTGAGTTG AAATATAGCC TCCTACTTAC ACCTGTTTCT TAGAAGCTGT  
30801 AGATATGAAG TAGCTGAAGC AGGCATTCCC TCTGAAACAT GTGTTTCACA  
30851 TATGTCATAA TTATCTTCTG CTCTCATTTT TCTTTTAGGC TTTTGTCTCC  
30901 ATCTCATTTT CCCTGTTTAC TCTCATTTTC ATATCTTTAC ATTTCTTTCT  
30951 CCAGAATTGT TCAGAAGCTT GGAACCCTTC ACTCCAGTTA TTCTTTGACT  
31001 ATGCAATTTG TTTCTGTGCT TCATGGCACT TATGGTTTGT AATCCTTGAC  
31051 TTGTTTGTAT AGCTCAGTGG TTAGGAGTAC AGTTTGGAGT TAGAATGCCT  
31101 GGGTTGAAAC TCTTAATTCT ACTCTACTTA CTAGTCTTGT GACTATAACA  
31151 AAATCTTAGT CCTCTCTTTG TCTGTAAAAT GGAGAGTATA GTAAATACAT  
31201 GGGCTTGTAT TAAGGATTA ATGAGTTAAC ATGTGAAATA CTTAGAACAA  
31251 TGCCTGGCAA ATGCTCAATG AATATTGAGT ATTGCTTGCT TTTGTTAGT  
31301 GCCATGCCCT TGTTCCTCAC TGAGGGCACA GACCATGTGT ATCTGGTTAA  
31351 CAGTTCTATG TCCACCACGT TGCAATAATG GACTCTCAGA AAATATTGAA  
31401 GAATATGTTA AAGAATGAGT AGAATTATGC TACTGAAAAG GGTGAGTGGA  
31451 AGGTAGGTAG GGGAAAGGAC ATATACAGCC CTGGAGGCAG CATATATGGG

FIGURE 3, page 10 of 42



31501 GAATGGGTCA CACAGTGTTC CTTGGTACTC TCTAGACCAT AGTGGGCCAC  
31551 CTCTTAGCTA GTGGCCTATG GATTATTTCA GCAGTCTGTT GGAAACATCC  
31601 ATGAATATGA TAATAATGAC CCATTTGTGG GTTCTAAGAA AAAGGACAAC  
31651 TACAATACTA GACAATAATA GTATGTAAGT TAGGAGGGAA GGGGATGATT  
31701 TGTATTAAC TGTCTAAAA TTCTTACCTT ATTTAGGATG ATGGGGTCAG  
31751 ACATTTAACTT TAGACTTTGT TATATATATG TGGTAAAATT TCAAGGTAAA  
31801 CCATTGAAAC TGTAAGTAGT GAGTATATAA CTTCCAAATC AGGGGGGAAA  
31851 GAAATGGAAT AAGAAAATAA ATACATAAAC ATAAGATTGA AACAATCCAA  
31901 TGAAGAGTAG AGAGAAGAGG GAAAAACATA GAAAGAATGA GATAATTAGA  
31951 AAGCAATAGG TAAGATGTGA GAAATAAAAT CAAGTACAGT AAAACTCCAC  
32001 TAAATGTGC CCTGCAGTAA TGTGGGGGCA TGATTTCCTT TCATCCCAT  
32051 TCTCAATGG GGCAGCCTAA ATAGCGTTCT TATCCTGTTT CCCTGGGGGT  
32101 TTGAGGTGGG TGACGAGTAA GTTAGAAGAT AATCACCTTC TGATCAGTTA  
32151 GGACTTTCTC AGTTTAGTCT TCAATTAATA AAAATTAATG TAAATTTTAT  
32201 CAGAAGGCAG AGATTGTCAG ATGAAAGAAC AAGCAAAATA AAAGTCTTAC  
32251 TGAAAAAAG CTGGGGTAGC TATGTTAATA TCAACTGTTA ATTATTATTA  
32301 ATAATCTATT AATAATAGAT TATATAGTAA AAACATTAAT AAAAATAGAG  
32351 TGTCACTACA TTTTAAAAAT CAGTATGAGG ATATACAATT TTTAAGCTGG  
32401 TTGATAAAAT TCTGGGGATT AATTGGCAAA TCCATCATAG TGGTGAGAGA  
32451 TTTTAAACACA ATTCTTCTCG TATTGATAG GTCAAGCAGA GAAAACTTT  
32501 AGTGAAGACA AAAACTTCTA AATACATAAG CTTGATTTAA TGGGCATGTA  
32551 ATAGGACCTA GCATCAAAAA ATTAGAAAAA ATATTTTTC TTAGGTATTT  
32601 ATGGAACATG TATAAAAAAT GATTTCGTAG TAGGCCATAA AGCCAGGTTC  
32651 AACACATTTT AAAGAAGTGG TATCACAAGA ACTGCTTTCT CTGACCACTA  
32701 TGCATTAATA TAGAAGTTAA TTACAGACAT AAATTATAAA AATGCCAATA  
32751 TTTTAAAGTG TGATATACAC TTCTCAACTT ATGGGTCAAA GGAAATCGTA  
32801 AGTGGAAATT CAAGGACACG TTGACTTGAA AACATTAAAA CTTATGGAAT  
32851 ATTTCTAAGA TGGAAGTTGT ATGAATTTTA TAGTCTGAAA GCTTTTATTA  
32901 GAAAAGAATT AAGTCTGAAA ATTAATGTGC TAAGTTAGGG GAGAGAAAAT  
32951 GGAATAATCT CGAAGAAGGT AGGAGGAAGG AGATAATAAA GAATATATAG  
33001 CAAAGATGCA GTAACAGGAT CAACAAAGCC AGAACTGTT GGAAAGACA  
33051 AGCCTCTGGA AAGATTGATG AAGAAAAAAG AGAAATGAGA TGTAAATAAA  
33101 TCATGTTTCTG TTATAAATAG GCACATAAGG ACTTTTAAAA AACTAATAAA  
33151 ATAATATGAA TCATTAATGC CAATAAATTT GAAACAGAC AAAGTAGGTG  
33201 AATTTCTAGA AAAATATAAC TTACTGGGAC TGAATGAAGA AGCAACAGCT  
33251 TATAGTACCT AAGCAATTGA AGAGATTGGG TCAGTAATTT AAAATTTTCT  
33301 CATAAACAAA ACGTTAGCCC CAGATGGTTC TTGCAAATGA TTAAAGAACA  
33351 GATGTACAAA CATTTCCAGA GTGTAGAAGT AACTGTCTCT ATCCTTTCTA  
33401 GGAGATCATT ATACACCAA AAGCAGACAG TATATGAAAC AGGGAAATTA  
33451 GAGGCCAAGA TACCTATGAC TTATATGTAA AAATTTAAAG AAAATATTAG  
33501 CAACTGAAT CAGCCATTTT AAAAAATATA CCACAATCAA TGCATTCATA  
33551 AGAGCAGCTT AACAAAATTT GTTAGAAGGC ATTAAGAAG ACTCAGTATA  
33601 GAAAAGATGT ACCTTCTCTC CAAATTGGTG ATAGAGATTC AATGCCATTA  
33651 AAAAAACCCA CCTGGTTTTT TTGAGGAAT TGTCAGCTG AGTCTCAAAT  
33701 TTATATCAAA GAGCAAGGC CTAAGAATAT CCAGGACATT CCTGAAGAAC  
33751 TGTAAGGAGC CAGGGGCTCG CCCTATCAGA TACCAAGGGT TGTATTAAAG  
33801 CCATAACCAA GTCAGTGCTG TTTCTACAGA AACAGACAAG TTAACAAGTG  
33851 AAACATAATA GAGAGCCCAG AAACAGACCC ATCCATATTT TGGATTTGTC  
33901 ACGTGAAGA AGTAGCTTTG CAAAACCTTG GGAAAGGAG AGTGTGTGCA  
33951 ATAGATGATG CTCGTGCTCA TGCAGACAAA AAGGAAATG GGATACCTGC  
34001 CTCTTACCCT ACACAAACAC CAACCTAAAC GTGAAAGTTA AACTATAACA  
34051 GCTTGAGGTG GTGGGGAAGA AATATCTTTA TCTCAGTGTA GGGAGAATTT  
34101 TATTTTAAAA AGAAGACACA AAAGGCCATA CATAGGAATG AAAAGATTGA  
34151 ATTCAGCTGC ATTAAGAAAG TTAATTCAG CTGCGTTAAA ATCAAGAGCA  
34201 TCTGTACTTG GACAGCATAG AGTGGAAAGA CAAAGAGAAG GTATTGCCA  
34251 GCTTATAACT TGAAGGATTA GAATGAATGA TATAAGAAC TATGTAATAA  
34301 AGAAAAAGAC ATACAACCGG TTAGAAAAAC GGGCAAGAG ATGAACAGCA  
34351 TATTTACCGT GAAGGAAACA GCGGTAGCAA ATGAACATGG TAAGAGATGC  
34401 TCAACACGTT TAGTAATTTG AAGGGAAATG CAAGTTATAC CCACAGCAAG  
34451 ACTATCTTAT CTAGGAAGTT TGTCAATACC CTAAATGTTT TGTGGTTTTA  
34501 AGCTACAGAG TTTGTAATTC ATTTATTTAT TCAATAAATA CTCAGTGGCA  
34551 GGCAGTCTT TAGAAACCTT GGTTATAACT TTGAATGAAA TTAAGAAAAA  
34601 TCCTTGCCCT GTGGAGGATG CTTATGTGTG GGGAGTTGGG TGGTGGGGTC

FIGURE 3, page 11 of 42

34651 AAACAACAAT TACATTAAAA TAGAAAATAG TGACATAAAT AAACCTATAA  
34701 ATATTGCAAC CCAGAGTTAT ATTATAAATG TAAGTAGTGA CTAGGACTCT  
34751 CATGCAGATA TACCTCTGTG CTGGGACAAA TGAAAGTTTA AGTGTAAATTT  
34801 CCCATATGCA AGTCAAAATA AAAAGTGACA CTAGAAAACA CAATAATGAA  
34851 TATCTGAAAA TTGCATTTTA TTTGACTGCC ATCCTTTTGC ATCATTTTCA  
34901 TACTAATTAT AGAATAAAAT TTGTAGGATG CACCAAAGCT TTTTITAGAG  
34951 ACATCCATTA ATTCAATAAA TAAATGAGCA CCTTCTTTGT GCCAGCAGCT  
35001 GTAAGAGGTG GCCCAAGGAA GGGAATAAAA CAGTCAAAAT CCTGGTACAC  
35051 TCAGAGTTTC TCTTAGGAGA AAACAGATAC AAATGGCATT AATTACCAAG  
35101 AAACCTGTAA AACAAGCCAA ATATTAATGA TAAATATTTG AGTACAGTAT  
35151 GTTAATTTTA AGATTGAAAA TGAGGTGCCA GGATTTCTTA AGACTCAAAG  
35201 GCGAAGATGG CTGAATAGGA ACAGCTCTGG TCTACAGCTC CCAGCGTGAG  
35251 CGACGCAGAA GACGCATGAT TGCTGCATTT CCATCTGAGG TACCGGGTTC  
35301 ATCTCACTAG GGAGTGCCAG ACAGTGGGCG CAGGTCAGTG GGTGTGTGCA  
35351 CCGTGCGCGA GCTGAAGCAG GGCGAGGCAT TGCCTCACTC GGAAGTGCA  
35401 AGGGGTGAGG GAGTTCCTTT TCCTAGTCAA AGAAAGGGGT GACAGATGGC  
35451 ACCTGGAAAA TCGGGTCACT CCCACCTGAA TACTGCACTT TTCTGACGGG  
35501 CTTAAAAAAT GGCGCACCAG GAGATTATAT CCTGCACCTG GCTCGGAGGG  
35551 TCCTACACCC ACGGAGTCTC GCTGATTGCT AGCACAGCAG TCTGAGATCA  
35601 AACTGCAAGG CGGCGGCGAG GCTGGGGGAG GGGCACCCGC CATTGCCCAG  
35651 GCTTGCTTAG GTAAACAAAG CAGCCGGGAA GCTCAAACCTG GGTGGAGCCC  
35701 ACCACAGCTC AAGGAGGCCT GCCTGCCTCT GTAGGCTCCA CCTCTGGGGG  
35751 CAGGGCACAG ACAAACAAAA AGACAGCAGT AACCTCTGCA GACTTAAATG  
35801 TCCCTGTCTG ACAGCTTTGA AGAGAGCAGT GGTTCCTCCA GCACGCAGCT  
35851 GGAGATCTGA GAACGGCGAG ACTGCCTCCT CAAGTGGGTC CCTGACCCCT  
35901 GACGCCCGAG CAGCCTAACT GGGAGGCACC CCCCAGCAGG GGCACACTGA  
35951 CACCTCACAC AGCCGGTTAC TCCAACAGAC CTGCAGCTGA GGGTCTGTCT  
36001 TGTTAGAAAG AAAACTAACA AACAGAAAGG ACATCCACAC CAAAAACCA  
36051 TCTGTACATC ACCATCATCA AAGACCAAAA GTAGATAAAA CCACAAAGAT  
36101 GGGGAAAAAA CAGAGCAGAA AAACCTGAAA CTCTAAAAAG CAGAGTGCCT  
36151 CTCCTCCTCC AAAGGAACGC TGTTCTCTAC CAGCAACGGA ACAAGCTGG  
36201 ATGGAGAAAT ACTCTGACGA GCTGAGAGAA GGCTTCAGAG GATCAAATTA  
36251 CTCTGAGCTA TGGGAGGACA TTCAAACCAA AGGCAAAGAA GTTGAAAACT  
36301 TTGAAAAAAA TGTAGAAGAA TGTATAACTA GAATAACCAA TACAGAGAAG  
36351 TGCTTAAAGG AGCTGATGGA GCTGAAAACC AAGGCTCGAG AACTACATGA  
36401 AGAATGCAGA AGCCTCAGGA GCTGATGCGA TCAACTGGAA GAAAGGGTAT  
36451 CAGCGATGGA AGATGAAATG AATGAAATGA AGCGAGAAGG GAAGTTTGA  
36501 GAAAAAAGAA TAAAAAGAAA CGAGCAAAGC CTCCAAGAAA TATGGGACTA  
36551 TGTGAAAAAG CAAATCTAT GTCTGATTGG TGTACCTGAA AGTGACGGGG  
36601 AGAATGGAAC CAAGTTGGAA AACACTCTGC AGGATATTAT CCAGGAGAAC  
36651 TTCCCCAATC TAGCAAGGCA GGCCAACATT CAGATTCAGG AAATACAGAG  
36701 AACGCCACAA AGATACTCCT TGAGAAGAGC AACTCCAAGA CACATAATTG  
36751 TCAGATTACAC CAAAGTTGAA ATGAAGGAAA AAATGTTAAG GGCAGCCAGA  
36801 GAGAAAGGTC GGGTTACCCT CAAATGGGAA CCCATCAGAC TAACAGCGGA  
36851 TCTCTTGGCA AAAACTCTAC AAACCAGAAG AGAGTGGGGG CCAATTTCA  
36901 ACATTCTTAA AGAAAAGAAT TTTCAACCCA GAATTTTATA TCCAGCCAAA  
36951 CTAAGCTTCA TAAGTGAAGG AGAAATAAAA TCCTTTTACAG ACAAGCAAAT  
37001 GCTGAGAGAT TTTGTACCA CCAGGCCTGC CCTAAAAGAG TTCCTGAAGG  
37051 AAGTGCTTAA CTTGGAAAGG AACAATCAGT ACCAGCCGCT GCAAATCAT  
37101 GCCAAATGT AAAGACCGTC GAGACTAGGA AGAACTGCA TTAACAAACG  
37151 AGCAAAATAA CCAGCTAACA TCATAATGAC AGGATCAAAT TCACACATAA  
37201 CAATATTAACT TTTAAATGTA AATGGACTAA ATGCTCCAAT TGAAAGACAC  
37251 AGACTGGCAA ATTTGGATACA GAGTCAAGAC CCATCAGTGT GCTGTATTAA  
37301 GGAAACCCAT CTCACATGTA GAGACACACA TAGGCTCAA ATAAAAGGAT  
37351 GGAGGAAGAT CTACCAAGCA AATGGAAAAC AAAAAAGAC AGGGGTTGCA  
37401 ATCCTAGTCT CTGATAAAAC AGACTTTAAA CCAACAAAAG TCAGAAGAGA  
37451 CAAAGAAGGC CATTACATAA TGGTAAAGGG ATCAATTCAA CAAGAAGAGC  
37501 TAACTATCCT AAATATATAT GCACCAATA CAGGAGCACC CAGATTCATA  
37551 AAGCAAGTCC TGAGTGACCT ACAAAGAGAC TTAAACTCCC ACACATTAAT  
37601 AATGGGAGAC TTTACACCCC CACTGTCAAC ATTAGACAGA CCAATGAGAC  
37651 AGAAAGTCAA CAAGGATACC CAGGAATTGA ACTCAGCTCT GCACCAAGCA  
37701 GACCTAATAC ACATCTACAG AACTCTGCAC CCCAAATCAA CAGAATATAC  
37751 ATTTTTTTTCA GCACCACACC ACGGCTATTC CAAATTGAC CACATACTTG

FIGURE 3, page 12 of 42

37801	GAAGTAAAGC	ACTCCTCACC	AAATGTAAAA	GAACAGAAAT	TATAGCAAAC
37851	TATCTCTCAG	ACCACAGTGC	AATCAAACCTA	GAACCTCAGGA	TTAAGAATCT
37901	CACTCAAAAC	CGCTCAACTA	CATGGAAACT	GAACAACCTG	CTCCTGAATG
37951	ACTACTGGGT	ACATAACGAA	ATGAAGGCAG	AAATAAAGAC	GCTCTTTGAA
38001	ACCAACAAGA	ACAAAGACAC	AACATACCAG	AATCTCTGGG	ACGCATTCAA
38051	AGCAGTGTGT	AGAGGGAAAT	TTATAGCACT	AAATGCCCAC	AAGAGAAAGC
38101	AGGAAAGATC	CAAAATTGAC	ACCCTAACAT	CACAATTAAA	AGAAGTAGAA
38151	AAGCAAGAGC	AAACACATTC	AAAAGCTAGC	AGAAGGCAAG	AAATAACTAA
38201	AATCAGAGCA	GAAGTGAAGG	AAATAGAGAC	ACAAAAAAC	CTTCAAAAA
38251	TTAATGAATC	CAGGAGCTGG	TTGTTTTTGA	AAGGATCAAC	AAAATTGATA
38301	GACCGCTAGC	AAGACTAATA	AAGAAAAAAA	GAGAGAAGAA	TCAAATAGAC
38351	ACAATAAAAA	ATGATAAAGG	GGATATCACC	ACCAATCCCA	CAGAAATACA
38401	AACTACCATC	AGAGAATACT	ACAAACACCT	CTATGCAAAT	AAACTAGAAA
38451	ATCTAGAAGA	AATGGATAAA	TTCTCGACA	CATACACCCT	CCCAAGACTA
38501	AACCAGGAAG	AAGTTGAATT	TCTGAATAGA	CCAATAACAG	GATCTGAAAT
38551	TGTGGCAATA	ATCAATAGCT	TACCAACCAA	AAAGAGTCCA	GGACCAGATG
38601	GATTACAGC	CGAATTCTAC	CAGAGGTACA	AGGAGGAACT	GGTACCATT
38651	CTTCTGAAAC	TATTCCAATC	AATAGAAAA	GAGGGAATCC	TCCCTAACTC
38701	ATTTTATGAG	GCCAGCATCA	TCCTGATACC	AAAGCCAGGC	AGAGACACAA
38751	CAAAAAAAGA	GAATTTTAGA	CCAATATCCT	TGATGAACAT	TGATGCAAAA
38801	ATCCTCAATA	AAATACTGGC	AAACTGAATC	CAGCAGCACA	TCAAAAAGCT
38851	TATCCACCAT	GATCAAGTGG	GCTTCATCCC	TGGGATGCAA	GGCTGGTTCA
38901	ATATACGCAA	ATCAGTAAAT	GTAATCCAGC	ATATAAACAG	AACCAAAGAC
38951	AAAAACCACA	TGATTATCTC	AATAGATGCA	GAAAAAGCCT	TTGACAAAAT
39001	TCAACAACAC	TTCATGCTAA	AAACTTTCAA	TAAATTAGGT	ATTGATGGGA
39051	TGTATCTCAA	AATAATAACA	GCTATCTATG	ACAAACCCAC	AGCCAATATC
39101	ATACTGACTG	GGTAAAAACT	GGAAGCATTC	CCTTTGAAAA	CTGGCACAAAG
39151	ACAGGGATGC	CCTCTCTCAC	CACTCCTATT	CGACATAGTG	TTGGAAGTTC
39201	TGGCCAGGGC	AGTTAGGCAG	GAGAAGGAAA	TAAAGGGTAT	TCAATTAGGA
39251	AAAGAGGAAG	TCAAATTGTC	CCTGTTTGCA	GACGACATGA	TTGTATATCT
39301	AGAAAACCCC	ATTGTCTCAG	CCCAAAATCT	CCTTAAGCTG	ATAAGCAACT
39351	TCAGCAAAGT	CTCAGGATAC	AAAATCAATG	TACAAAAATC	ACAAGCATTC
39401	TTATACACCA	GCAACAGACA	GAGAGCCAAA	TCATGAGTGA	ACTCCCGTTC
39451	ACAATTGCTA	CAAAGAGAAT	AAAATACCTA	GGAATCCAAC	TTACAAGGGA
39501	TGTGAAGGAC	CTCTTCAAGG	AGAAGTCAA	ACCACTGCTT	AATGAAATAA
39551	AAGAGGATAC	AAACAAATGG	AAGAACATTC	CATGCTCATG	GGTAGGAAGA
39601	ATCAGTATCG	TGAAAATGGC	CATACTGCCC	AAGGCAATTT	ACAGATTCAA
39651	TGCCATCCCC	ATCAAGCTAC	CAATGACTTT	CTTCACAGAA	TTGGAAAAAA
39701	CTACTTTAAA	GTTCATATGG	AACCAAAAAA	GAGCCCGCAT	TGCCAAGTCA
39751	ATCCTAAGCC	AAAAGAACAA	AGCTGGAGGC	ATCATGCTAC	CTGACTTCAA
39801	ACTATACTAC	AAGGCTACAG	TAACCAAACC	AGCATGGTAC	TGGTACCAAA
39851	ACAGAGATAT	AGACCAATGG	AACAGAACAG	AGCCCTCAGA	AATAACGCCG
39901	CACATCTACA	ACTATCTGAT	CTTTGACAAA	CCTGAGAAAA	ACAAGCAATG
39951	GGGAAGGAT	TCCCTATTTA	ATAAATGGTG	CTGGGAAAAAC	TGGCTAGCCA
40001	TATGTAGAAA	GCTGAAACTG	GATCCCTTCC	TTACACCTTA	TACAAAAATC
40051	AATTCAAGAT	GGATTAAAGA	CTTAAACGTT	AGACCTAAAA	CCATAAAACC
40101	CCTAGAAGAA	AACCTAGGCA	TTACCATTCA	GGACATAGGC	ATGGGCAAGG
40151	ACTTCATGTC	TAAAACACCA	AAAGCAATGG	CAACAAAAGC	CAAAATTGAC
40201	AAATGGGATC	TAATTTAACT	AAAGAGCTTC	TGCACAGCAA	AAGAAACTAC
40251	TATCAGAGTG	AACAGGCAAC	CTCCAAAATG	GGAGAAAAAT	TTTGCAACCT
40301	ACTCATCTGA	CAAAGGGCTA	ATATCCAGAA	TCTACAATGA	ACTCAAACAA
40351	ATTTACAAGA	AAAAAAACAA	ACAACCTTAT	CAAAAAGTGG	GTGAAGGACA
40401	TGAACAGACA	CTTCTCGAAA	GAAGACATTT	ATGCAGCCAA	AAAACACATG
40451	AAAAAATGCT	CACCATCACT	GGCCATCAGA	GAAATGCAAA	TCAAAACCAC
40501	AATGAGATAC	CATCTCACAC	CAGTTAGAAT	GGCAATCATT	AAAAAGTCAG
40551	GAAACAACAG	GTGCTGGAGA	GGATGTGGAG	AAATAGGAAC	ACTTTTACAC
40601	TGTTGGTGGG	ACTGTAAACT	AGTTCAACCC	TTGTGGAAGT	CAGTGTGGCA
40651	ATTCTCTCAGG	GATCTAGAAC	TAGAAATATC	ATTTGACCCTA	GCCATCCCAT
40701	TACTGGGTAT	ATACCCAAAG	GACTATAAAT	CATGCTGCTA	TAAAGACACA
40751	TGCACATGTA	TGTTTATTGT	GGCACTATTC	ACAATAGCAA	AGACTTGGAA
40801	CCAAGCCAAA	TGTCCAACAA	TGATAGACTG	GATTAAGAAA	ATGTGGCACA
40851	TTTACACCAT	GGAATACTAT	GCAGCCATAA	AAGATGAGTT	CATGTCTTTT
40901	GTAGGGACAT	GGATGAAATT	GGAATCATC	ATTCTCAGTA	AACTATCACA

FIGURE 3, page 13 of 42

```

40951 AGAACAAAA ACCAAACACC GCATATTCTC ACTCATAGGT GGAATTGAA
41001 CAGTGAGAAC ACATGGACAC AGGAAGGGGA ACATCACACT CTGGGGACTG
41051 TTGTGGGGTG GGGGGAGGGG GAGGGATGCG ATTTGGGAGAT ATACCTAATG
41101 CTAGATGACG AGTTAGTGGG TGCAGCGCAC CAGCAAGGCA CATGTATACA
41151 TATGTAACAT ACCTGCACAT TGTGCACATG TACCCTAAAA CTTAAAGTAT
41201 AATAATAAAA AAAAAAGACT CAAAGGCACA GTCACAGACA GTTTGATTTT
41251 TTATAATAGC TGTTAATTTT CCTAACTTCG AGGAAGTTGA TAGCATGTTT
41301 TGAGTATATT TCAAACTAC ATTCAAATGT TGCAATAGAA CATTAAAGAAT
41351 TATCTTCATG ATCCACTAAG TGCATGAAAA AAATGGATAA TGAATCTATT
41401 CATTACCATC GTTTAATATT TTATCTTCAA GTTTTTGTGT TTTGTAGCTC
41451 ATTTGGCAGAG TTTGACAGAG TGCTGAAAGT ATTCTTTAGT GAGCTGGCTG
41501 TAATTTTTTG GCCCATTTTT ATCTAGATAA TTAAACTAT CTGACAGGAC
41551 CATAAAATGC TTTGCTGCCAT TTCCAACAAC CTATATTGTG GGATGGGGTT
41601 TTTTAATTTA ATGAGAATAT TATGTTAGAA AAGAACTGT CATTCTGTAA
41651 AGTGGCCAAT AATGTTAGTT TTATTTATCA ATTTAGTTTT GTACTTTGAT
41701 CATTTTTTTA AAATTTTCAGC ATTGATGTTG ATGGGACAAT GACAGTGGAC
41751 TGGAATGAAT GGAGAGACTA CTTCTTATTT AATCCTGTTA CAGACATTGA
41801 GGAAATTATC CGTTTCTGGA AACATTCTAC AGTAAGTCTA CTTTATGTAT
41851 TTATACTTAT TTGGAGCTAT AAACCATAGG TACAGTTATC ACCCAAGAAC
41901 ACTCTGTAAC ACTTATGGGC CAGGATACCT GAGTCCAGT AGCTCCTTAA
41951 CCTGTAGAGT TCTATTTATT CTATTAGGCA TAGATTTATA GAGTATTAAA
42001 CAAAAAATAA CAGCTCTCCC TCTCCCTCTC CCTCTCTCTC CCCCTCCCA
42051 CGGTCTCCCT CTCCCTCTCT TTCCACGGTC TCCCTCTGAT GCCGAGCCAA
42101 AGCTGGACTG TACTGCTGCC ATCTCGGCTC ACTGCAACCT CCCTGCCTGA
42151 TTCTCCTGCC TCAGCCTGCC GAGTGCCTGC GATTGCAGGC GCGCACCACC
42201 ACGCCTGACT GTTTTTTCGTA TTTTTTTGGT GGAGACGGGG TTTTCGCTATG
42251 TTGGCCGGGC TGGTCTCCAG CTCCTGACCG CGAGTGATCC ACCAGCCTCG
42301 GCCTCCCGAG GTGCTGGGAT TGCAGACGGA GTCTCGTTCA CTCAGTGCTC
42351 AATGGTGCCC AGGCTGGGGT GCAGTGGCAT GATCTCGGCT CGCTACAACC
42401 TCCACCTCCC AGCCGCTGTC CTTGGCCTCC CAAAGTGCCA AGATTGCAGC
42451 CTCTGCCCCG CCGCCACCCC GTCTGGGAAG TGAGGAGCGT CTCTGCCTGG
42501 CCGCCATCG TCTGGGATAT CAGGAGCCCC TCTGCCTGGC TGCCAGTCT
42551 GGAAAGTGAG GAGTGTCTCT GCCCGGCCGC CATCCTGTCT AGGAAGTGAG
42601 CGTCTCTGCC CGGCCGCCCA TCGTCTGGGA TGTGAGGAGC CCCTCTGCCT
42651 GGCTGCCCAG TCTGGAAAGT GAGGAGCGCC TCTTCCCGGC CGCCATCCCA
42701 TCTAGGAAGT GAGGAGCGTC TCTGCCCGGC CGCCCATCGT CTGAGATGTG
42751 GGGAGCGCCT CTGCCCCGCC GCCCGCTCTG GGATGTGAGG AGCGCCTCTG
42801 CTCGCGCCGC CCGTCTGAGA AGTGAGGAGA CCCTCCGCCG GGCAGCCGCC
42851 CCGTCTGGGA AGTGAGGAGC GTCTCCGCCG GGCAGCCACC CTGTCCGGGA
42901 GGGAGGTGGA GGGGTACGCC CCCC GCCCGG CCAGCCACCC CATCCGGGAG
42951 GTGAGGGGTG CCTCTGCCCG GCCGCCCTA CAGGGAAGTG AGGAGCCCTT
43001 CTGCCCCGCC ACCACCCCAT CTGGGAGGTG TACCCAACAG CTCATTGAGA
43051 ACGGGCCATG ATGACAATGG CGGTTTTGTG GAATAGAAAA AGGGGAGAGG
43101 TGGGGAAGAG ATTGAGAAAT CGGATGGTTG CTGTGTCTGT GTAGAAAGAG
43151 GTAGACATGG GAGACTTTTC ATTTTGTCT GTACTAAGAA AAATTTCTCT
43201 GCCTTGGGAT CAGTTTGATC TATGACCTTA CCCCCAACCC TGTGCTCTCT
43251 GAAACATGTG CTGTGTCCAC TCAGGGTTAA ATGGATTAAG GGCGGTGCAA
43301 GATGTGCTTT GCTAAACAGA TGCTTGAAGG CAGCAGGCTC GTTAAGAGTC
43351 ATCACCACCT CCTAATCTCA AGTACCCAGG GACACAAACA CTGCGGAAGG
43401 CCGCAGGGTC CTCTGCCTAG GAAAACCAGA GACCTTTGTT CACTTGTTTA
43451 TCTGCTGACC TTCCCTCCAC TATTGTCTCTG TGACCTGACC AAATCCCCCT
43501 CTGCGAGAAA CACCAAGAA TGATCAATTA AAAAAAAAAA AAAAAAACA
43551 ACCCAAGACT GCATAAATGT CCATTCTGAA AACTTGGAAG AAGTACCACC
43601 TTGATGAATA AGCTGTCTAG CTTTTATTGG CATTTAAGTA TTCTGCCATA
43651 GGAAGTGTA AAAGTTGTAG GCTTTTACTT TTTATAGGTA CTATATTGTC
43701 CAAATAATCT CAGCACCTCA TGGTTGCTAA GGATCTGTGT CCTTGTTTGG
43751 TCAGATTATG TTTATCTCTG GCATAAGGCA CTTAACAATA TTCATTAAAG
43801 GTTACAGAAAT CTTTTTGCTT CATCTGCTTA GCATTTCTA CCAGTTTGTT
43851 TTCCACCAAA CTTTCAAATT TTGATTGTTT CATTAAATTT CTGCATACTG
43901 ATGTAAACCA AGTTCTATTA TTGTGCAATC TGCTCCTGAA ACCCTTAGGA
43951 ACTCTCTGAA GGAGTTTTAT TTATTTTTTG TTTTTGTTTT TGTTTTTGTT
44001 TTGTTTTTTT GAGACGGAGT CTTGCTCTGT TGCCAGGCT AGAGTGCAGT
44051 GGTGCGATCT CGGCTCTCTG CAAACTCGGC CTCCGGGGTT CACGCCATTC

```

FIGURE 3, page 14 of 42

44101 TCCTGCCTCA GCCACCGGAG TAGCTGGGAC TACAGGCACC CACCACTGCG  
44151 CCTGGCTAAT TTTTTTTTGT TTTTGTAGTAG AGACGGGGTT TCACCGTGTT  
44201 AGCCAGGATG GTCTCGATCT CCTGACCTTG TAATCCGCCC GCCTCGCCTC  
44251 CCAAAGTGCT GGGATTACAG GCGTGAGCCA CTGTGCCCCG CCTTTTTTTT  
44301 TTTTTTTTCT TTATGGGCTT GTCTTCTACA CTTGAGATTT GACTAAATTA  
44351 AATATGCATT AAATGAAGTC AGGAGTTCAC ATTGCCACTA GTAACAATGC  
44401 CTAAGCTTAC ATAAAGCATT ATAAAATTGT TGGTGATTAG TGCCTTCTCA  
44451 GCTATGAGTA TAAGATAATA TTATACTAGT AGTTCAGTTG CCTAGATAAA  
44501 TTGTACACTA TGTGAAGTTT TATTTACATA ATTCTTACGG TATTTTAA  
44551 GGTAGTTGAT AACAGTTGAG ACTACAATTG TATCTCCATT TTATTGATAG  
44601 TAAAATGAAG GAAGGGAGGG TTACTIONTACC AGGAGAGCTC CTCCCCGTTG  
44651 CACTCTTGGC TGTAAAAAAT TTTCTGCCAA AACAATTAG ATAATAGAAT  
44701 TGTAAAAATA TTATTATAGA ATTGTTTCTC TCAAACTATA GTAATGTAGA  
44751 ATAGGTTGAA GGGGTGATGA TTTGAAACAA TACCTCTCCA TTAGCTAAAT  
44801 TTTATATAGA ATCTATTGCA TGTTTTAAAT GATAAGTCAG ATTTATAAAA  
44851 ATATTTTTTAT AAACAGTAGG AAATGAGTTT AGGGGTATTC ACATACAGTT  
44901 TTAATTTTTA TTTACATATT TAAAACATAT CATGGTATAA ATATGATGTG  
44951 GATATAAAT TGAGATAAAG GAAGTATTGT TTAAGAATTG ATGAACTAAT  
45001 TTCTTAAAG ATGTCATCAC CAGTTGGTTT TCTAGCCTTA TGAAAAATGG  
45051 TTGCAATAAA AAAGATTGAC TATGATAAAA TGCTGCCCTT TCATTTTAAC  
45101 CTAGACCAAG AGAAACATA CTGTGAATCT ATGATGAATG AAAGAAAGTT  
45151 GTAAGTGTG GTTTTGTATA TTTGTAATTA CTGTTTATTT TCATTCTTGT  
45201 TGAAGTGTG CTGTACTTTG TTCATTGTGA GTAGACAACT TATAATCTAT  
45251 GTACTCAAAT TGGTTTAGTA TAAATCTAG GGAATGAAGT TCATATTAAC  
45301 TGTAATAATA CATGATTGTT CTCTAAACAA AAACGTCTTC TGGGATTATT  
45351 TTTAACTAAG GCGCATGGGG ATCTTTTTTT CATTTTACAA GGAATTGAC  
45401 ATAGGGGATA GCTTAACTAT TCCAGATGAA TTCACGGAAG ACGAAAAAAA  
45451 ATCCGGACAA TGGTGGAGGC AGCTTTTGGC AGGAGGCATT GCTGGTGCTG  
45501 TCTCTCGAAC AAGCACTGCC CCTTTGGACC GTCTGAAAAT CATGATGCAG  
45551 GTGAGCTTTA TTATCGTGTG TCCAGGTTTG CCCTAAATAT TCTAAAAACA  
45601 TGAGAAATGT GGTGCTTTGA AAAAGAAGTT TTAATAATTC TCAGTAATAA  
45651 TCTTTTATAC CCTAAAAAAT AAATCTATTT TGTGTCTGTT AACTCTAAAT  
45701 TCAGTCCATG TAAGTATGGC AGTGTACCAA ACCTTAAATT GTTAGTACAT  
45751 GTGTGTAATG AACTTTTAAT CTTTGGCATT CTATGACTAT TCAAACATTT  
45801 AATTCAAAAA ATATCTCTAG CTATTGTTGT AGGATTCTCC TGATTATAG  
45851 TTTCTTCTT TTTAATATAC TTTATCAAAA GTAAAGTATT TTTGAAATCT  
45901 AGACTCTTAG AGCAGCAATG TAATTTTGAA AATTATTCTA AAGCTGAGGT  
45951 TAGCAGAAAA AGACTGGCT TTATAGACTG ACTTTGCTAT TTAGTAGCAG  
46001 TGTAGCATTG GGCTGGCCAG AGTGGAAAGA GGAATGGAA AAGAATTAAT  
46051 ATGTATTTGC TCACTGTGGT AACCCAGTTA ATCCTTGCAG CAGCCAGTG  
46101 AAGTAGGTAT TTTATCATTT TTCCAGGGGG AATCTGAGGC CCAGAGAATT  
46151 GACTTTTCCT TTACAACAAA TGAGAGGGGG AATGCAGTAT CTTTGCCTCC  
46201 AGTGCTCCTG GTTCTCATGC TGCATGAAAC CTCTGAGGTC TCATTTTCCT  
46251 TCATTCTGGG ATGGGGATAA GAATATCTAA TAAGAATGGT TTAAGAATCA  
46301 AGCAATATCA GGTATGTGAT AATGCTGGT AACTGGAAT AACCTATTGG  
46351 AACATAGTAG TTGTTTACAA AATATTTTAA AACTTTGTT ATACTTATGG  
46401 TCAACACTTT TTATATTTGT CTGTAGATTT CTGTACAAAA AGATTCTGAC  
46451 ACTGTTTTTA GCCAGCATTC CTTCAGAATG TACCCAAATC TCAAAATTTA  
46501 TTTAGGGGCA AAGCTAATGC TTTAAAGAAA AAGGAGAGGG GATTGGTGTG  
46551 TGTTTTTCTT TAGGAACAGT AGTAACTTGA CTTTATAGAG ACTTGAATAA  
46601 GCATTTATTT TTTCTTTTGT CCTATTTTAT TGTGAAGTTT ATTTATTTAA  
46651 AATAAAATGG ATTTCTCTGG AATTTAGTTT CTGCAAAATT GAGGAGTTTC  
46701 CAAAGTCAAC CTTCAGGTTT GATACTTCTC TAGAAAGACT CACATAACTC  
46751 ACTGAAAGCT TATTACCCCT GGTTATGGTT TATTACGGGG AAAAGATGCG  
46801 GATGAAATC AGTCAAGTAA AGAAGCACAT AGGGCAGAGC TTCTGTGTGTC  
46851 CTCTCCCTGT GGAGTCTCCA TGTCTTACTT TCCTGGCACT GTTATGTGGC  
46901 ACTAGGCATG GAATATTGCA GACCAACCAG GGAAGCTCAC CTGAGCCTTT  
46951 GGTGTGCAGA GTTCTTATTG GGGCCTGTTT TCATACTGGC CACATGGCTG  
47001 GCCTTCAGAA TTCAACCCGT TCTGTGAGTG TGTGTGTGTG TGTGTGTGTG  
47051 TGTGTGTGTG TGTTTAGTGG TAGTCACCCC TTTTATGTGA GCTGAAACAA  
47101 TCAGAAGAAT AGCTGATTTG TTTAATTATT TTTGGTGTAT TGGACTTAAT  
47151 CAGTTTTTAT CTGTAGGTGG TCATAAGGTA CAGTATTTT AAGTACTAC  
47201 CACATCTGTA GTATAAGCCA AGTAATTTAT CAGTACTCAC AGGATGGGTA

FIGURE 3, page 15 of 42

47251 CATGTTGTAA TGAATTTATT GCCTAGAGAG GGCCTCAAAA TATGCCAAAG  
47301 AGGGTGCAAT TTTTATTTTT GGTTTCAGGC TGTATGCATT CCAGTGTGG  
47351 TAGCCCTGAT ATACACAATA TCCAAACCAT TTCAGACCCA TTTACAGTTC  
47401 ATGCTGTGAC TACTTCTTGA GGAGAGGGAG TAACATATTA CTTTAAATTA  
47451 TATGTAATAA TATACATACA TTAAATTATA TGTAATAATA TAATATTATT  
47501 ATTTGCAGTA TACTTTTTTA TTTCCCTTTA ACTGAGCTTG TTCATGTTTC  
47551 AAAGGGGTGT CCATTGCCTG ATACATAATT TAGTTAATAT TATCTTATGA  
47601 AGGTTGTTC TAATTTTAAAT ACTCTTCTTG TCTTCTCTCT CTGCTTCTC  
47651 ACACTGAAGA TACCAATTAT TCTTAGTTT AGAGTCAGAG ACAGGCCTCT  
47701 AAAATCATGG CAATACTCCC TCTCATCATT ATATATATTT TTCAACCTTT  
47751 CTATATTTTA TTTTCAAATA TATCTTCTTG CAGTTAGAAA CGGTATTGAA  
47801 AAAGATTGTG TGGTTGTTCT AGAAAAAGTA ATAGTAATAT GCCACCAGCA  
47851 TTTTATATCA TTCTGCTTTT ATTTTATAGG TCACGGTCA AAATCAGACA  
47901 AAATGAACAT ATTTGGTGGC TTTCGACAGA TGGTAAAAGA AGGAGGTATC  
47951 CGCTCGCTTT GGAGGGGAAA TGGTACAAAC GTCATCAAAA TTGCTCCTGA  
48001 GACAGCTGTT AAATCTCTGG CATATGAACA GGTAATTGTT ATCACCCGTG  
48051 GAATTTATTA ACAAAGAGGA GTTAGTAAAC GGATTCAATA AATGTTAATG  
48101 TATAATGCTT TTGGGATTCT TGTTTTAATA CATGATAATC TTTCACATAT  
48151 ACCCATTAAG GAGGATCACT TATAGGAGAT TAGACTAAAT AAAATCAGAG  
48201 ATTTCTCATG ACCAAGTTAT GGGATTCTTA ATTCATCATA TTATTTATAA  
48251 AGTTTTTTTT TTCTAAGTAG TTCTTAAAGG AAGGGTAGAA TTTTAGTTTA  
48301 TTCATTCTGA ATCCTGAGCA GAAGCAGCAC ACTAACATAA GTTTTATGAA  
48351 AGTGTCACAA TCTAACCTCT GGAAGGAAAA CTATAAGTTG AAGTCCTTTG  
48401 TGTAATTTGA CGTTGCTGTA AAATTGAGCT GAGTTTGGAG TGACACCTCC  
48451 ATGAAGGCAG GGGCGTGGCT TCTTCCCCAT GTACTCCAGC ACCTAGACAG  
48501 AGCTTGGCAT GTGATAAGTT TCAAGCGAGT GTTGAATGAG TCAATGAATG  
48551 AACAAATGCA TTTACCTCTG AATCACTTCT CTGTCGGCTT TTGTAACTT  
48601 GGATTATTTG AGCTATTGCT TCAGCCTAAC TCAATGTAAA GGGGAAATAC  
48651 AGAGGTAGT TTTAGAGTTT GGGTTCTCTT TATGGTCATT AGCAGAACTG  
48701 TCTAGTTGAG CAGCCACAGA TTATGTTTTC CATTATTTAT TCCATCATTG  
48751 TTTTCAAGG ACTGTAAGGG CTTGAAATT CAACTCCCCC CCCCATAGTT  
48801 TTTGTATTAT TCCATGTAGA TTTTAGATTA TTCTGGAGAG TGTTTTGTTC  
48851 TTGAGCAACA GAATACTCTT GAGAAGATTA CGAAGTCCAG TGGTATCCTT  
48901 TTCTTTGCCT AGGAAATAGA GAAGCAAAAA AAAAAAAAAA AAAAAATTAA  
48951 AGAAAATCTA GTCTCCAGGA TTTTAATTAG AACCTATCCT TGGGAAGGCT  
49001 ATTTTCCTTA TATGAAGGTT TGAAGATTCA AATCATGATT ATTAAGGGCT  
49051 AATGTTTGAG ATACCCTTAG GTTATTCTGA CCACATACTT GGATTTTATG  
49101 ATAGGAAAGC CACAGCCTAA AATAAATAAA TACTCAATGC AGTTATTTCA  
49151 GTATGCAAGA AGTTTGGTAT TTTTGAAAAA GTCCATGGGT ATTGCAAGCA  
49201 AATATGCACA TTTTGCTTTA TGCCATTGTG CAGATTCTTA CCTTGGATAC  
49251 CACCAACAGG CATCCTCTGC TTCTGTCCAC CCAAGCTCCT TCCTGAGACC  
49301 TCTTTATAGT ATTGTGATTT CTGCACACTA ACTTTCTTAG ACATGAAGAG  
49351 AAAGCTGTCT ACACAGTGTG GTGTAGTTTT CTTATGGGCT CTGGACCTAT  
49401 GGTGCTGTTT TCTCTCCTCC TGCTGAAGGT CCATTTCATCC CTCGGGGCTC  
49451 TCTAAAAGCC ACCTTCTGTG GACAAGCATA TACTAAGCAT CTCAATCAAA  
49501 GCCAGTTCCCT CCCCTGTCCA GCCTCCCTCG AGTGCTGAAT TGCAGAATAT  
49551 CCCATTTTTC ATTGGATGAT GGAAAACCCA TTGTTTTCCC AGTGGATTGT  
49601 AAATTACTTC GGGGTAAATA GGCTGTATAT ATTCTCAAAT TTCCAGAGT  
49651 ATGTAAC TAGTACTTTA GGAACTAAG AAAAAGATCT TTCAACCTG GTATGTAGCT  
49701 TAGTACTTTA GGAACCTAAG AAAAAGATCT TTCAACCTG GTATGTAGCT  
49751 CTGTCAAACA CATCATCAGT ATGGGGTAAA CCTGTGTTCT CTGTGGGTTG  
49801 TCATTACCAT AGTAGTGTCA TTGTATCATT GACAGTGTA TAGTGTGGG  
49851 TAGTGTCTT GTGGTTTCAG CTGCCACTCT GTACTGACTG CTTTCCACTC  
49901 CAACATCTTC CTCTTTATCT CAACACTGTA GGTCTACCTG TGTACTGTGT  
49951 GTTTCAGCAT CTCTGCTTGC ATGACCCAGG AGTGCCTCCC ACTCAATATG  
50001 GCCACCATGC ATGGTCATCT TTCTGCTACT CCCTGTCTCC TGACCTGCT  
50051 CCAGCAACAC AGACAGACAC CCTTCCTCTT TCTATATGTC ATATGGTGGG  
50101 GAATGCCCTT TAGTACTTAC TCAGGAGTTA GTTCCTCTGG GAAGCCTTCT  
50151 GTTCTAGTTT CTTTTGTTA CAGCACTTTC ACATTGAATT CTGACGTTCT  
50201 CTGTACTTAT CTGCTTTGTG AGACTGTGAG CTTCTTAGG CAGTAGCTAC  
50251 TTGTATTCTT AGCACCTTGC CCAGTGCCAG GAAACCCTTA TTAAGTAAAT  
50301 GAAAAGACAG AACTGACAGA CTGGAATTAG AGCTCAAGCT TGCCTCAATC  
50351 TCAAGCCATT AAGATGAAGG GGAGCCGGGC GTGGTGGCTC ACGCCTCTAA

FIGURE 3, page 16 of 42

50401 TCCCAGCACT TTAGGAGGTA GTTTGCTTGA GCCCAGGAGT TCAAGACCAG  
50451 CCTGGGCAAC GTGGCAAAAC CCCATTCTTA CAAAAAATAT AAAAATTAGT  
50501 TGGACGTGGG GGTGTGTGCC TGTACTCAGG ATGCTGAGGT GGGAGGATCA  
50551 CTTGAGCTCG AGAGGCAGAG GTTGCAGTGA GCTGGGATCA CACCATTGCA  
50601 ATCTAGCCTG GGTGATAGAA TGAGACCTTG TCTCAAAAAA AAAATAAATA  
50651 AATAAATAAA GGGGAAGATA AGGATTGGAA ACAGAAGGAG CAGCATGTGG  
50701 ACAGAAATGT AGGCACAAGA AGGCATCACT CACTGAAGAG ACTGAAAGTG  
50751 GTTCACTGTG CCTCAAGACT GGTGGAGTGT GTTCCGGAA AGATAATGAT  
50801 GAAAGAGCTG GACAGATAAA CAGGGGCCAA ATGTAATAGG AGTCTGGATT  
50851 TTATTCTGAA TATGGTAGGG GCTATTGTAG CATCTTATAT AGGGAAGTGA  
50901 AATGAGTACA TTCACATTTA AGGAATATCA ACCTGAAAAA AGAGTGGAGA  
50951 CATTGTGGG GGTGATGAG GAGGTGTGAG GTAGACTAGA GGCAGGGAGA ATATTTAAAT  
51001 AATTGAGGTA AGAAATGATG AACACCAGTA TAAGGTGATG TCTTTAAGGA  
51051 ATGGAGAAGG GAATGAACTG AGAAATATTT TGGAAGTAGA ATCAACAGAA  
51101 CTCCTGACT GACTGGATAT GGAGGTGAGA AAGAGAAGAG TCAAGAATGA  
51151 TATTCTAATT TCTAACTTGA GTGACTGCAT TCAAAGAGAA TACAATATCA  
51201 GGTTCATTT TGTGCATGCT GAGTTTGAGA TGTGTGGGAC ATGTACAGGG  
51251 AGCTGTCCAG TAAGCAATTG GGTATATCAG CTAGCCATTA AGAGAGAGAT  
51301 CTTTGTATAGA GAGGTGTGTTG CTGAGTTGAG CCATTGGAAT GGGCAGGATC  
51351 ACTCAAGAAG AGCTTATAAA TGAGAAGAAT TCTAGGAATA AGTCCAAAGG  
51401 GAGAAGTAAA AGAAGAACT TGCAAAGGAC ACTGAGAAGA AATAGCTCGA  
51451 GGGATGGGAG AAAATCCAGA GAGAGGGATG GCATAGGAGT CAGTGAAGG  
51501 AAACGGTTTC ATGGGGGTCA GTACTACTGG GTAGTGAATA TAATAAGAAT  
51551 ATCTTTTAGG ATTTCTCAAC CCAGAGATAG GTAAGCTTAG TATAAATGCT  
51601 TCTGTGAAGT AATGAAATGA GAAACCATGC TGAAATGAGC TTAAAGTGAA  
51651 TGGGAGGTGA AGAACTTGG ACAGTAGAGA CACATTTTGA GGGAGTTTGA  
51701 CAGTGAAGAG AAGGAACTA GAAGAGGGAG AGGGTGATAG ATAAGAAAGA  
51751 TGTGGGTGG AGGGGATTTG TTTTTTTGTT TTTTGTGTTT TTTTCTGTTT  
51801 GTATGTTTGT TTGTTTTTGA GATGGAGTCT CACTTTATCA CCCAGGCTGG  
51851 AGTAAAGTGG TGCAATCTCA TCTCACTGCA ACCTCTGCCT CCTAGGTTCA  
51901 AGTGATTCTT CTGCCTCAAC CTCCTGAGTA GTTNNNNNNN NNNNNNNNN  
51951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN  
52001 NNNNNNNNNN NNNNNNNNNN NNNNNNTGCCT CAGCCTCCCG AAATGCTGGG  
52051 ATTGCAGGAG TGAGCCCCC GTGCCTGGCC TGGAGGGAGG ATTTTGATTT  
52101 GACTTTAATG TGCCTGTTGC TGAAGGAAGC ATGTCAATAC AAATAAAGAA  
52151 GTTGAAACA TAGGTAAGAG AGGTTGATTA ACCCGGTAGG TGTTTCAAGG  
52201 GAGTTTGTGT GTAGGGAAAG GGAGTGGGAG ATGGAAGGG GCTGGGGGAG  
52251 ACAGGTCTTA TCCAGAGACT GTTAAAAGGA TTAGTCTTTG ATTACAAGAA  
52301 GAACCTTCT TATACGTGTT TGGGAAGAAA AAATATGTGA GTAGCTATGG  
52351 ATAATTTTGC AGGAGGTGGG CAGAATACCA AGATATTCTG CCTGGTGGCC  
52401 TCTCTACTCT TCCTTGAGCT CCTGAGAAAG GATGTGATCT GAGAATGAGG  
52451 GAGGAAGTGG TATTGGAAGC TGGAGGAGAA TGGAGAAGAT CAAAATGGTT  
52501 AGTCTAACAA ATGGGAGAGA ACTGAGATAG ACAAAGGAT TTCAGGGTGG  
52551 TTTTGAGGGC TCAGTTAAGT CTCCTTTAGG AAGGTTCACT TCTGTAGCCT  
52601 TGGCAAGTTA CTTAAAGTCT CTGTGACTAT TACCTCATCT CTAAGATGGG  
52651 GACTAAGCTT GGTGACATAG TTTTACATAC CAGGCACAGT GCCTGACTTT  
52701 TTGGCTCTGT CCTGAAGTCT TCCCTTTGTA TATGGTATGT TTCGGGGAAT  
52751 AGGAGCCTCA AGCACTTATC CTTTAAATAT TTATCTCCA TCAGTCACTA  
52801 AACGTTTACT CTGTACTTTT GATAGGTGCT GTGGGGGTCC AGGGTATAAA  
52851 AGGTACCTTC AAAGTTACTG TTAAAGTGCA GGAAGGTTTT TAAGCAAATT  
52901 ATGTTTAAATG ATTTTGACAA TCTGACATGC AGGAAAATTA ATAGGCGCTA  
52951 TGCAGAAGAG GAGTTTATG TAACACTCTG TAGTTCAGGA AACAGAGCCC  
53001 TTGGAAGCAG TGATCTCTCT GGGGAGGAAT GTCTGGTATT TGGGAATCTC  
53051 ATGAAATGAT AATATACTTA ATTTTATCA TGAGCAGCAA AACACAGATT  
53101 TGCTAGGAGA AAGTCATCGT ATGTTGTTGC ATTGGGCACT TTAGATCCCA  
53151 GGGAACAGAA ACTGGCTGGC ACAGGAATGG GCATCACTGT GGGGATGGAT  
53201 CATGTAGGGG AAGGATCCCT GGAGAAGTCC AGGAGGTGAG ACTTCCCCCT  
53251 TCCCTTCTCC ATGCATGAGT CCACTTCTCT CTGTTGACTT TCCCCTTGTC  
53301 CCTCTGTTGA CAGCAGCTGC TTACCTCTGG AGACCCCTC ACATTCTGA  
53351 GAGAAGGAAT CTGGCTTGCC TGGCTAATTC CCATGGTCTA TGTGTTGGGCA  
53401 GAATGTCTTA GCAAGTTGTG TAAAGATAGT GTATTATAT ATTAATAATA  
53451 ATAATAACAT CTACTGAACA TTTGCTAGGT GTTCAGACCT GCACTAACCG  
53501 TGTTACAAGT ATTATTTTTT TGTAATCCTT TCCATAACCC TGTGAGGTAA

FIGURE 3, page 17 of 42

53551	GTACTGTTAT	CACAGACAAG	GAAACCACAA	TGTGGACCTG	TTCATGAACT
53601	TGCTCGAGGC	CACGTGGCTC	TGGAGTTCCA	GCTCAGGTCT	GCCTGACTCT
53651	CAATCCCATG	ATATTAATAT	ACTGGCCAGT	CACTATTTTG	GCTGTATTGG
53701	GGTCATATTT	ATACCCTTGG	TCCAGTTAGC	TATGTTGGGT	CACTTTAGTA
53751	CTGATAGCCA	GGGAGATGCT	GGGCTTGATA	GGTTAGTATA	ATTCTATGTA
53801	TTACCTACAA	AAACTGTTT	TATAAATTGT	TTTGTTAACA	TTTGTTTGTC
53851	ACCTATTTAT	TCATTTTATT	TGCACTGGTG	AAAATAAACT	CATCTTTTAA
53901	AAACTGTGGG	GAAAATATCC	AAACATTGTG	AAAACCTGAT	TAACCTTGTA
53951	TTTTCTGTAC	ACCTGGGGAG	GGATGCTGTT	ATGCTGTTTC	AGCAAAGGAG
54001	CAACTTGGTC	CAATCTGGGA	GACATCTGTG	TTTTGTGGAA	ATCTGACTTG
54051	AAAACCACTG	TCCAGTCACT	GCGTGTATTA	GCATTTAGGC	CTTGCTCTTC
54101	TGCTATGTAT	TATTAATGTA	GTGTATACAT	TTCGAGACAC	ATCATCACAT
54151	TTGTCAATTT	ATTGATTCT	AGGAGCTGAT	TTGTATTCTA	GGATTGTCTA
54201	GTTGGCTTGG	GCTGCCATAA	AATACCACAG	TGTGTGTGGA	ATCAACAACG
54251	GAAATTTATT	TCTAACAGTT	TCAGAGGCGG	GAAAGCCTAA	GATCAAGGGC
54301	CAAGCCAGTT	TGATTTCTAG	TGAGCGTTCT	CTTCTCAGCT	TGTAGACAGC
54351	TGGTATGTGC	TCACATGGTC	TTTTCTTGGT	GCACATGTGA	AGGGGGAGAG
54401	AGAGAGTGGG	CTCTCTGGTG	TCTGCTCTTA	CAAGAACACT	GATCCTGTCA
54451	TGAGGGCTCC	ATCCTCATGA	CCTCATAACC	CTAATTACCT	CCAGAAGCCT
54501	CATCTCCTAA	TACCATCACA	TGGGAGGTTA	CAGCTTCAAC	ATATGAATTT
54551	GGTGGGGGTG	CAGCTCAGTC	CACAGCAGGT	AGTAATGTGC	ATTTTAAAC
54601	TTGTTTATAC	AGTACAAGAA	GTTACTTACT	GAAGAAGGAC	AAAAAATAGG
54651	AACATTTGAG	AGATTTATTT	CTGGTTCCAT	GGCTGGAGCA	ACTGCACAGA
54701	CTTTTATATA	TCCAATGGAG	GTGAGTACCA	TTGTCAAGTC	TGACTGTGTG
54751	ATGGTGTTCG	TGTTGGTTGT	CTATTGCTCT	CTAACAAGTT	ATCCCAAAT
54801	TAACAGTTTA	AAACAAGCAT	TTATCATCGC	ACAGTTTCTC	TGGGTCAGGA
54851	ATCTGGAAGC	AGCTTAGCTG	GGTGCCTCTG	GCTCAGGGTT	TTTCACAGCC
54901	CACAGTCAAG	ATGGTAGTCA	GAGCTTGGA	TCAGCTGGAG	GCGGATTCCA
54951	AGCTCACTCA	TGTTGCTGCC	AGGCCTCACT	GGCTATTGGC	TGGAAACATC
55001	AGTTCCCTTAT	CACGTGAGCC	TTTCTGTAGG	CTGCCTGAGT	ATCCTCAAAA
55051	CACAGTAGCT	GGCTTCCCTA	GAGTCAGTGG	TCCAACAGAG	AGAGAGAGAG
55101	AGAGTGCCCTA	AGATAAAAGC	TGGTATCTTT	TGCCTCTTCT	GCTGTATTCC
55151	ATTGATCACA	CAGACCAACC	CTGGTAGAGT	GTAGGAGGGG	CTGGTATAAT
55201	GGTGTTAATA	ACCGGAGACA	AATATCACTG	GGGGTCACTT	TAGAGGCTGG
55251	CTGCCACTTT	AGAGGCTGGC	TGCCATTCCCT	GTCCAAAGAG	TTTCTGTACC
55301	ATAAATTTAA	TAATGGAATC	TCAGGATTTG	ATTATATGGT	GATTATCCTA
55351	ATTAGACATC	CTTTCATTAG	TGCATAGGTT	GGCAAAACAC	AGACCTACGG
55401	ACTGTTTCAT	ACAGCCCTTG	ACCTAAGAAT	GCCTTTTACA	TTTTTAAAAA
55451	GTGGGCAACA	CAGGAAAAAG	TGAGAAAGAT	CTAAATCGA	CACCCTAAGA
55501	TCACAATTAA	AAGAACTAGA	GAAGCAAGAG	CAAACAAATT	CAAAAGATAG
55551	CGGAAGACAA	GAAGTAGCTA	AGGTCAAGAG	AGAACTGAAG	GAGATAGAGA
55601	CACGAAAAAC	CCTTCCAAAA	ATCATTGAAT	CCAGGAGCTG	TTTTTATGAA
55651	AAGTTTAAACA	AAATAGACAA	CTAGCCAGAA	TAATAAAGAA	GAAACCAGAG
55701	GAGAATCAAA	TAGCCCCAAT	AAAAAATGAT	AAAGGGGATA	TCACCACCAA
55751	TCCCACAGAA	ATACAAACTA	CCATCAGGGA	ATACTATAAA	CACCTCTATG
55801	CAAATAAACT	AGAAAATCTA	GAAGAAATGG	ATAAATTCCT	GGACACATAC
55851	ACGCTCCCAA	GACTAAATCA	GGAAGAAGCT	GAATCCCTGT	ATAGACCAAT
55901	AACATGTTCT	GAAATTGAGG	CAGTAATTAA	TAGCCTACCA	ACCAAAAAAA
55951	ACCCAGGACC	AGACAGATTC	ATAGCCGAAT	TCTACCAGAG	GTACAAAGAG
56001	GAGCTGATGC	CATTCCCTTCT	GAAATTATTC	AAACAATAGA	AAAAGAGAGA
56051	TTCTCCTCTA	ACTCATTTTA	TGAGGGCAGC	ATCATTTCTGA	TACTAAAACC
56101	TGGCAGAGAC	ACAACCAAAA	TAGAAAATTT	CAGGCCAATA	TCCCTGATGA
56151	ACATCAATGT	GAAAATCCTC	AATAAAATAC	TGGCAAACTG	AATGCAGCAG
56201	GACATCCAAA	AGTTTATCCA	CCATGATCAA	GTTGGCTTCA	TCCCTGGGAT
56251	GCAAGGCTGT	TCAACATATG	CAAATCAATA	TAACGGAATT	CATCAATAAA
56301	CAGAACCAGT	GACAAAAACC	GCATGATTAT	CTCAATAGAT	GCAGAAAAGG
56351	CCTTCGATAA	AATTCAACAC	CACTTCATGT	TAAAACTCT	CACTAAACTA
56401	GTTATTGATG	GAATGTATAA	CAAAATAATA	AGAGCTGTTT	ATGACAAACC
56451	CACAGCCAAT	ATCATACTGA	ATGGGCCAAA	GCTGGAAGCA	TTCCCTTTGA
56501	AAACCGGCAC	AAGACAAGGA	TGTCCTCTGT	CAGCACTCCT	ATTCAACGTA
56551	GTATTGGAAG	TTCTGGCCAA	GGCAATCAGG	CAGGAGAAAG	AAATAAAGCG
56601	TATTCAGATA	GGAAAAGAGG	AAGTCAAATT	GTCTCTGTTT	GCAGTTGACA
56651	TGATTGTATA	TTTAGAAAAC	CTCCTTGCTC	CAGCCCCAAA	TCTCCTTAAG

FIGURE 3, page 18 of 42



56701 CTGATAAGCA ACTTAAAGCA AAGTCTCAGG GTACAAAATC AATGTGCAAA  
56751 AATCACTAGC ATTCCCTATTA ACCAATAATA CACAAACAGA GAGCCAAATC  
56801 ACGAGTGAAC TCCCATCCAC AATTGCTACA AAGAGAATAA AATACCTCGG  
56851 AATACAACCT ACAAGGGATG TGAAGGACCT GTTCAAGGAG AACTACAAAC  
56901 CACTCCTCAA GGAAATAAGA GAGGACACAA ACAAATGGAA AAACATTTC  
56951 TGCTCATGGA TAGGAAGAAT CAATATCATA TCATAGGAAG AATCAGTGGC  
57001 CATACTGCCC AAAGTAATTT ATAGATTCAA TGATATCCCC ATCAAGCTAA  
57051 CATTGAATTT CTTACAGAA ATAGAAAAAA CTACCTTAAA TTTCATATGA  
57101 AACTAAAAAA GAGCCTGTAT AGCCAAGACA ATCCTAAGCA AAATGAACGA  
57151 AGCTGGAGGC ATCACGCTAC CTGACTTCAA ACATACTACA AGGCTACAGT  
57201 AACCAAAACA GCATGGTACT GGTACCAAAC AGATATATAG ACCAATGGAA  
57251 CAGACGCAAA GCCTCAGAAA TAACACCACA CGTCTACAAC CATCTGATCT  
57301 TTGACAAAAA CAAGCAATGG GGAAAGGATT CCTTATTAAA TGTATGGTGT  
57351 TGGGAAAACCT GGCTAGCCAT ATGCAGAAAA CTGAAACTGG ACCCCTTCCT  
57401 TACACCTTAT AAAAAAAAAA TTAAGTCAAG ATAGATTAAA GTCTTAAACA  
57451 TAGACTTAAA CTATAAAATC CCTAGAAAAA AACCAGAGCA ATACCATTC  
57501 GGACACAGGC ATGGACAAAG ACTTCATGAC TGAATCACAA AAGCAATGGC  
57551 AACAAAAGCC AAAAATTGACA AATGGGATCT AATTAACTA AAGATCTTCT  
57601 GCACAGCAAA AGAACTATC ATCAGAGTGA ACCGGCAACC TACAGAATGG  
57651 GAGAAAAATT TTGCAATCTA TCCATCTGAC AAAGGGCTAA TATCCAGAAT  
57701 CTATAAGGAA CTTAAGCAAA TTTACAAGAA AAAAAAACC ACCAAAAAGT  
57751 GGGTGACGGA TATGAACAGA CACTTCTCAT AAGAAGACAT TTATGCAGCC  
57801 AACAAACGTG AGAAAAGGCT CATCATCCCT GGTGTGTAGA GAAATGCAAA  
57851 TCAAAACCCC AATGGCATA CACTCTCACGC CAGTTAGTTA AAAAGTCAGG  
57901 AAACAACAGA TGCTGGCAAA TATGTGGAGA AATAGGAATG CTTTACACT  
57951 GTTGGTGGGA GTGTAAATTA GTTCAAGCAT TGTGGAAGAC AGTGTGGCAA  
58001 TTCCTCAAGG ATCTAGAACC AGAAATACCG TTTGACCCAG CAATCCCATT  
58051 GCTGGTTATA TACTCAAAGG ATTATAGATT TTTCTACTAT AAAGACACAT  
58101 GCACACGTAT ATTTATGCA GCACTGTCA CAATAGCAA GACTTGAAC  
58151 CAACCCAAAT GCCCATCAGT GATAGACTAG ATAAACAAAA TATGGCACAT  
58201 ATACACCATG GAATACTATG CAGCCATAAA CAAGGATGAG TTCATGTCCT  
58251 TTGTAGGAC ATGGATGAAG CTGGAAGCCA TCATTCTCAG CAACCTAACA  
58301 CAGGAACAGA AAACCAACA CCACATGTTT TCACTCATAA GTTGGAGTTG  
58351 AACAATGAGA ATACATGGAC ACAGGGAGGG GAACATCACA CACTGGGGCC  
58401 TTTTGGGGA TGAGGGGCTA GGGGAGGAAT AGCATTAGAA GAAATACCTA  
58451 ATGTAGGTGA CAGGTTGATG GGTGCAGCAA ACCACCATGG CACGTGTATA  
58501 CCTATGTAA ACACCTGCAC GTTCTGCACA TGTATCCCAG AACTTAAAGT  
58551 ACAATTTTAA AAAAGTAGGC AAAAACAAAA GAAAAGAAAA GTAATATACA  
58601 ACCGAGAAGT ATGATTTTAA GCTTGCACAG ACAGATATTT TACTATTTAG  
58651 TCTTTACAGG AAAAGTTTTC CAACTACTGC TTTATAGCAA AAATAATATT  
58701 GTAGATGTGG AATTTATTGA TATAGCAGAG GGGTTTTTAT TAAGTATGA  
58751 CTTAAGCAAG ATAAATACAA TTTTCACCGA TATGTGGTAT GCATGCTAAT  
58801 ACAGCTTTTT TTAAGCATCT TAATATGATT GTTTATATTA CTCCACACAC  
58851 CTCTCAAAAA AACTTAATAC CCTATTTTTC CTCTCATATC CTCCCATATC  
58901 AGTTAATAGT ATCACCTTCC CAACTCCCCA CTGCCCCATC CTGTGTTCAC  
58951 AGCTAGAAGT ATTGGGGTTA TCCTTTATAC TACCATTTC CTCACCTTCC  
59001 AGATGCAGGT GGTCACCACT CAGTTTGTGT AAGACATCAA TAGATTATCT  
59051 TGCTTCCATT TCCTTGGTCA CTTCTTCAT CAGATCCTCC TTGCAGTAAA  
59101 CGGGTCTCTC TGGCTTTGGT CTTAGCCCC CAATAGAGGT AATACATGAA  
59151 AGAGAATGTA TCAACAAAT GTACAGTCTT TTGAGTGACA ATATGTGCTA  
59201 GGTATTTGTT CCATGTAAAA TTAATTCATT TGAATCCCAT GATGATAGAG  
59251 TTAATATGAA CAATCATATT TTGTTTTTTT TTATATCCAG GTTATGAAAA  
59301 CCAGGCTGGC TGTAGGCAAA ACTGGGCAGT ACTCTGGAAT ATATGATGT  
59351 GCCAAGAAGA TTTTGAACA TGAAGGCTTG GGAGCTTTTT ACAAAGGCTA  
59401 TGTTCCCAAT TTATTAGGTA TCATACCTTA TGCAGGCATA GATCTTGCTG  
59451 TGTATGAGGT GAGTTTGTAG AAATCTTTTG AATTGGAAAA TGCAGTTAGA  
59501 TCTTGTTAGA ATTGGACTTT ATATGAAGAA GTAGATATAT ACCAGAAAAC  
59551 AGTGTGTGAC CAGAAGTAAA TTCAAGCATG TGTATTATAG ACTTTCAAGT  
59601 AACTTGAGTG TGAATATGCA TGGGGTCACT TTTGTATTAG ATTTTCTTGG  
59651 GAATTGCTTT TGTTAATGAA GAGTAGACTC AAAGTTAGGT ATAGTTGTTC  
59701 ACCTTAAAG GTGTTTCTAG AGATTTTTC CTTTGTTTTG GATTTGCAAA  
59751 AATCTGACAT TAAGCCAAGT GACTAATGTG ACTAACATGA GTAATACAGT  
59801 TTCATTCCTT GTACGGAAGA ATACAAATCT TGGATCAACC CTGCAATCTA

FIGURE 3, page 19 of 42

59851 AATCATTTAA TAATTTATGA ATCTCACAAA CAATTATTGA GCACACACTA  
59901 TACAAACCAC TAGGTTAGAC ACTGGATCTG GGGATTCAAA GGACTCAATG  
59951 TGTGCCTTGA AGAAACTGAA GGTCTGGTGG GGGAGACAAA CGACTAAAAC  
60001 TCAGCGTGGT TATCTGTGCT GCGACAGACA TGAGCCAGGG TGCATGTTAG  
60051 GATGAGACCT AAGCTACAGC GTAGAGGAAG AGTGGAAATGT GTAATGAAAA  
60101 GAAGAGTCGA ATTTTTTTTT TAAAGAGCTT TATTGAGATT TAGTTCATAT  
60151 TCCTTACATT TCACTCATTT GAAGTGTACA AGCAAATGGT TTTTGGCTTC  
60201 TTACATAATT TTTAAAAAAT ATTATAAAAT ATAAAAATTT CCATTTTACT  
60251 AATTTTAAAGT GTACAATTCA GTGGCATTAA TTACATTCAC AATATTGTGC  
60301 AACCATCAAC ACTATTTCCA AATCCTTTTC CTCACTCCAA ACAGAAACAC  
60351 CTTAACCTTT AAGCAATAAC TTCTACCCT CCGTAACTCA AACCTTTGGT  
60401 AACCTCTAAT CTGCTTTCTA TGTCTAGGAA TTTACCATT CAAGATATCT  
60451 TATAAGTAGA ATCATACAGT ATTTTTCTTT TTGTGTCTGA TTTATTACTC  
60501 TTAGCATAAT GTCTCTAAGG TTTGTTTCATG TTGTAGCATG TATCAGAACT  
60551 TCATTTCTTT TCATGGCTGA GTAATATTCC GTTATGTGTA TATACCACAT  
60601 TTTGTTTTAGT CCTTCATCTG TTGAAGAGCA TTTGGATTAT TTCTACTTTT  
60651 CCAACATTGT GAATAATGCT GCAGTGAACA TTGGCATCTG CGTATCTGTT  
60701 CGAGTCTATG CTTTCAATTC CTTTGGGTAT ATATCTCAGA ATGGAATTGC  
60751 TGAGCCATAT GGTCAATTCTG TGTTTAGCTT TTAGGAAC TAAGACTGTT  
60801 TTCCATAGTG GCTGCACTTA CATTCTCACC AGCAACATAC AAAGGTTCAC  
60851 GTTTTTCCAC GTCCTTATTA AACTTAATT TCCATTTTAA AAAAGCTTAT  
60901 TTTTATTATG GCCGTCCTCT TAGGTGTGAG GTGGTATGGT TCAGGACTTT  
60951 ACTTCTTGTG CTGAGTTTTT TAAAAAATG TGATTAAAAA CACATAACAT  
61001 AAAGTTTATG ATTTTAAACA TTTTAAATA TATAGTACAG TAAGTGTTAA  
61051 CTGTTTGTGG TTTGTTGTGC AACAGATCTC TAGAACTTTT TCACTTCTCA  
61101 AAACCTTAAAC TCTATAGTCA TTAAACAACA GCTCCCAATT TCCCCTTCAC  
61151 CCCAGCGCTG TGTAACCTAC TTTCTCGTTT TATGAGTTTG ACTACATTAA  
61201 ATACCTTGTA TAAGTGAAAT CATGTGGTAT TTCTCTTCC GTGACTGGCT  
61251 TATTTTCATG AACATAGTTT CCTCATGATT CATCCATATG ATAGCATACA  
61301 ACAGGACTTT TTTGTTTTTA AGGCTGAATA ATAATTTGTT GGGTATATAT  
61351 ATCATAATTT CTTTATTCAT CTGTTGATGG ACATTTGGAT TGTTTCTACA  
61401 TCTTGACTAT TGTGAATAGT GCTGCAGTGA ACATGGTTGT GCAAAATATCT  
61451 CTTCAAGATA CTGTTTTTCAG TTCTTTTTGA CATATACTCA GAAGTGGAAAT  
61501 TTCTGGGTCA AATGGTAATT CTATTTTTAA GTTTTGTAGG AACCTCCATG  
61551 TCATTTTCCA TAGTAACTAG ACCTTTTTGT TTTTAAACAT TTCTATCAAT  
61601 GTACACCAAG ATTCCAATTT CTCCATGTCC TCCCAACAC CATTAAGTGG  
61651 GGTGGTGGTC TACTACTATT GCTGTGTTGC TGTTTATTCC TCCCTTCAGT  
61701 TCTGTAAGTG TTTGCTTCAT ATATTTAGGA GCTTAATATT AGGTCCATAT  
61751 GAAGTTATAA TTTCTTCTG GTAAAGTGAC CCATTTATCA TTATGTAATG  
61801 TCCATCTTTG TCTCTTGTGA CAGTTTGTGT CTTAAAATCT ATTTTGTCTG  
61851 ATGTAATTAT GGCCACCCCT TTTCTCTTTG GGTTCCCGTT TTTATGGAAT  
61901 ATCTTTTTC ATCCTTTCAC TTTCAGCTTA TGTGTGTCC TAGATCTAAA  
61951 GTGAGTCTCA TAGATAAGGT ATAGTTGATT CTGTATGTGT TATTCACTCA  
62001 GCAATTTATA TCTTTTAGTT AGGGGATTTA ATCCATTTAT ATTTAAAGCA  
62051 GTTACTGATA GGAAGGACT TACTGTTGTC ATTTGGCTAG CTACCTTTT  
62101 ATCTTTGTCC TGTGGCTTTT CTGTTTTTCC CTTCCTCTCT TCCTGGCTTC  
62151 TTCTGTGTTT TGTTGATTTT TTTTTTTTTT GTAGTGATAT GTTCTGATTC  
62201 CCTTCTCAT TCCCTTTGTG TGCATTCTAT AGATGTATT TTTGTGGTTA  
62251 CCATTGCAAC TACATAAAGC ATACTAAAGT TATAGCAACT TATTTTAAAGC  
62301 TGTTTACAAC TTAACCTCAG TGGTATATAA AACTCTATTT CTTTACATAT  
62351 TTCACCTCCT CCCACAAC TTTATGTCTT TTGATATTGT ATATCCTTAA  
62401 CATAGATTTA TAGTTACTTT TTATGCTTTT CTTCCTTAAA TTCTGTTTAA  
62451 ATTTTGTTTT TGAAATTTAG ATTTTCAAGT TATTTATATA CCTTCATTAC  
62501 AATACTATAG GATTTTATAA TATTCTAAAT ATTGACCTTT ACCATAGAGT  
62551 TTCATATTTT GTGGTTTTGT GTTGCTATTT ATCATCCTTT TGTTTCTCCT  
62601 TTTAGCCTTT CTTGTAGGGC CGGTCTAGTG GTGATAAGCT GTATCAGCTT  
62651 TTGTTTGTC GGGACAGTCT TAATTTCTCC TTTTTTGAAG GGCAGTTTGT  
62701 CCCATACAGT ATTTTGTGTT GGCAGTTT TTAAGTTTCA AAACATAGAA  
62751 TATAACATTC CATTTCTTTC TAACCTGCAA GATTTCATT GAGAAATGCA  
62801 CTCAATGGAT TTTTAAATCC ATTGAGATAA TTTTAAATC CTGTAGGATT  
62851 TAAAATTTT AGTCTTACAG GATTAAAAA TTAAGGTT AAACCTGTTA  
62901 TATAACATAT TAACATGTAT TTTATACTTA AAGTATCTTA TGTTTAAAAA  
62951 GTTGATTATC ATATATATTT TATACAGTTT CTCCTAATTA TTGCCTTCTA

FIGURE 3, page 20 of 42

63001 ATGAAATACA GGGACCTAGA GTAACAGGGA TAAAGTATGG CCTTTTGATC  
63051 AGCACGCCTG GTTCTGAGTC CTTCTTAAAA AAACCTCTGGG CCTGGTGTGG  
63101 TGGCTCATGC CTATAATCTC AGCACTTTGG GAGGCCGAGG CGGGCGGATC  
63151 ACCTGAGGTC AGGAGTTTGA GATCAGCCTT GCCAGCATGG TGAACCCTG  
63201 TCTCTACTAA CAGTACAAAG ATTAGCTGGG CGTGGTGGTG GGTGCCTGTA  
63251 ATCCAAGCTA CTCAGGAGGC TGAGGCAGAA GAATCGTTTG AACCTGGGAG  
63301 GCAGAGATTG GGCCACTGCA CTACAGCCTG GGTGACAAGA GCGAGACTCC  
63351 ATCTCAAAAA AACAAACAAA AACTCCGCTG AGATGAATTT TTCTCATTTT  
63401 TAAAATCAGA ATAATAGATT TATGTAAGAG TTTCTGTAAG GCTCAAATGA  
63451 AATATATGTA ACGTGTAAAA TGAGATACAA TTAGTAGAAT TATATTATTT  
63501 TATTAATACT CACCATAAGA GGTGTTCTTT AGATCCTGCA GCGTTTGCTG  
63551 CGCAGTTTAC GTTTGTTTAG AAGAATGTCA GTAACCGGTG CAAACCTCAT  
63601 GTGTTCCGCA CCCCCAGTGG CCTCCCACCT CTCCACAGAG TCACCGCCTC  
63651 CTGCAGTGCC TGCTGCTTCT GCAAATGCGT GGCCTCATCC TGCAGAAACG  
63701 GGGCTTCTCA TGAGGTTGAG AATAGCTGTG AAAATGTTTA CGTTGAAGTT  
63751 GTAGAGTTCT TTAATTATTT TCTTCTTTAT TTCTCTGGCA GCTCTGAAG  
63801 TCCTATTGGC TGGATAATTT TGCAAAAGAT TCTGTAAACC CTGGAGTCAT  
63851 GGTGTTGCTG GGATGCGGTG CTTATCCAG CACCTGTGGT CAGCTGGCCA  
63901 GCTACCCATT GGCTTTGGTG AGAACTCGCA TGCAGGCTCA AGGTGAATTT  
63951 TTGATTACAG AACCACACCG ATAAAAGTGC TGCACCAGTA ATGTGCTTTT  
64001 AGAACTCCAA GTTCTACTAA GATGCAGACT GTAGTTTAA GACAGTATTT  
64051 CTCAACCTTT TTTTCATTAT TGCCCTCCTA AGGAATCTTT TCAGAAATTC  
64101 TTTTCTAAAA TGCTCCCTCG TCATGAAATT TTAATGCGAC AGAAGCATTG  
64151 CATATGTACT GTATGCATAC ATATGCCTTA TAGATAAACA GAGTACTATT  
64201 TTTTTTGACT GTGTTACATG CACGTTTAA GATTATAAGC TTTAGTATCT  
64251 GATGGATTG GGTTCAGATC CTTGCCTCAG ACTTCTGGG GTTTTAAATG  
64301 GGAATGAAAA TTGTACAGTG TTGTAAGAAT TACCAACAAT ATAAATAAAG  
64351 CATCTTGGGT TTGTTAAATT TTTGGTAAAT GGTGGTTGGA ATCATTTTTT  
64401 AGTGTTGCGT AGACCCACCA AGTTTTGAGC TGTGATTCCT CCTCACTGTG  
64451 ACACGTGCTC CATTGTTGGC TTTGATTACA CTGTACCATC CTGGTTGTTC  
64501 TGCCAGCCCA TTGATAACTT TTACCATTG CTGGCTTTTA TTGCTATCCC  
64551 CACTCTATTA AAGTATGCAT TCAAATGCCT TTCTTTTCTC TTTGATGCTT  
64601 TCCCTGGTCA GTCTTATCCA TTGTTTTCTT AAGTAGTACA CCTTGGGCAT  
64651 CTACAGCTCT ATTCCCAACC TCCCTTCCAA GTGCCAGCCA CAGCAACCCC  
64701 AGCCAAGCAG TCAGTAACTA ATTTGGCAAAT ACTCCCTGAG CCATTGTCCC  
64751 ATTCTAGACA CTGCCAGATG CTAGGGGTAG AGCAGTCAAC AAGTCAGGTG  
64801 TGGCCCCGCC AGTGTAGAGT AGAGAAGACG TTATGTCCAG CAAGTAAACA  
64851 ACCTGGTTAA ACCAACTCCT CTTTTGTTAG GGGAGCACAG AGCAAGGAGC  
64901 TATAACCTAA AATGGGCGCT GCAGAAATGCT GTCAGTGAAG CTGAGACTGG  
64951 AAAGATGAGT GGGAGTTAGC TGGGCACAGG CCAGTGGAGT GGGAACAGAA  
65001 AACATTCCAG TTGAGGGAAA GCATGTGTGA AGACACTGAG GCAGGCACCA  
65051 ACATGGTGTA TTTAAGGAGC TGAGAGACAG TCATGGCTGT AGAGAAAAAC  
65101 ACAAAGTAGT GAACTACACG TTTCTTGTGT ATTCTCTCAT TTCACCATCA  
65151 TAACCATCTT GGGGATGGGA ATACTAACAT TATCCCCATT TTTCAGATGA  
65201 GCAACTGGGG CAGAGAGAAT TTAAGTAACT CCCACAAGAT TATACCTGTG  
65251 GTAAATAGTG GGACTGAAAT TCAGACACAT GCAGTCTGAT TCTAACCCTC  
65301 CTGTCTGCCA GCTCTGATCC AGAACTTTGC ATGACTGATA CGGCTGATAG  
65351 ATTGTCTATG GCTGATAGAC TGTCATTTCT GACCTAAAAG TCTGATCATT  
65401 TTACATCTGT TCAGACATCT TTGCAGCCTT TCGGTGTCAG TTCCAAAGTT  
65451 GTTAGTGGGA ATTTCAAAGC CTTTAATAAT CTAGCCCCAC TTTGTTCACT  
65501 CTCTGTGTAA TAACCACATA CAACAATTGG CTGCATCTCC ATAGCACATG  
65551 GTACTCCTCC CGTTGTCTTG GTTGTGCCAG CAACACTGGT TTTTCGCTTC  
65601 TCTTCCTGCT TGTTGAGGTC ATTTCCAAGG CCCAGGTCTT TGTGCTTTTT  
65651 CCCAAGCTTC CCAGAGCTTC TTCCATACTC CCCTTACTTC CTGAGATTTA  
65701 ACTGTTCTCT CTTCAGCGCT TGCTAGTAA GAAGGAGGCA GCAGCAGCAC  
65751 TGTGGGGTGG TGGAAGTGT ACCAGCTTTG GAGTCAGACC ATTGATCTC  
65801 AGCCCTACCA TTTTCTACTT AGATTTTTTT AGGACAAATT TCTCCATCTT  
65851 TCTAAGCCTC CAATTGCTCA CTTACAAAAT TGATATAACA TTTACCTTGC  
65901 AAGATTGGTA TGGAAAGTAA TTAACCCAGT ATTTAGAACA TAGTAATTAA  
65951 TAAATAACTA TTATTACCAT CATTACTATA GTTAGGACAC TCACTGTTAG  
66001 GTGCTATACA AAGAGGATCA TAAAAGGGAT GTTGTCTTGG GCTTCTTGGA  
66051 ATAAATGTTG TCCTTTTACT GTATTTTAGA ATATCATTTT GGGTCATAAT  
66101 TGTTTGTTGT CATAATAATG AAACATACTT GAATATTTAA TTACCTCTTT

FIGURE 3, page 21 of 42

66151 TTTTATTTT TTAGCCATGT TAGAAGGTTT CCCACAGCTG AATATGGTTG  
66201 GCCTCTTTTCG ACGAATTATT TCCAAAGAAG GAATACCAGG ACTTTACAGA  
66251 GGCATCACCC CAAACTTCAT GAAGGTGCTC CCTGCTGTAG GCATCAGTTA  
66301 TGTGGTTTAT GAAAATATGA AGCAAACTTT AGGAGTAACC CAGAAATGAT  
66351 GTTGCAATTT TTGCTTTAGC CTGATAATTG AAACCTTCAA CAATCTCTGG  
66401 AGTGACTTTT TCTCCTCGAA TTGAAACAAG TCTATGGCAA AAGAAGCTGC  
66451 ATTTTTTTTCA CAAAAGGGAA GATGGTAACA ATGGTCACTT CAAACTTTTG  
66501 GGCTAAATTA TATGTACACA GAAATGTTCA AAATCATAGT TTTAATGTGT  
66551 TTTGAAAAGG CCACACAATT ATACTTTATC TTTTCTTAAT AATCCTGCAA  
66601 ATCTCTGCCC TGAATCCGAA ATCTGAAAAT GTACTGGCTT GAACAAAATT  
66651 TGTTTTGTGT GTTAGAGTTA TAAATCATT AATCTTTATT CGGGTGGTTT  
66701 ACGTTTATGC CAGTTCCCTT ATATTTAAAT TTCTTGTTT ATATATTTTG  
66751 AATGTCTTTA TAGATTTCTT TAAATTTCTT TATAGAACCA TTAATAGAAA  
66801 ATCATTACAT TTAATATATA CCTTACAGCA AAAGCATCCA AATAAGTATA  
66851 GGGTTTATGT CCTTATTTT CTTTCAGCTG AATACGAATG AGCACAGTGG  
66901 TGGAATTTCT GAAGGGAAGT GATGAAATTA TATTTATTTC AGTGGGCACT  
66951 TTTCCATTTT ACCACTGTAC CATTATTTGG TTCCTGGAGT TATACACTAA  
67001 TTTTCAGTAT ATTACTGTTA AATTACCAAC ACAAGGCAAT TTATTTGAAA  
67051 GATTCCGTTT ATCTGCCAT TGCTTTGAAA AGCAGCAGGA AACGAAATCC  
67101 TTTGACTTGT ATCAGCTTCT GCAGAGCATC TTTGTTTTC TTTGTCCTTT  
67151 GTTTCCTACC TTTTGAATCA GATTCCGTTT TAGTCAGGAA GACTTCTTGG  
67201 GACCATTCTT AGTAACCTGA AATTTCTTTT TTAATTGCAT GAAGTGGATT  
67251 GATCATGAGC AAATGATGTG CTTATTTCTC CCTCACTGTT GAATATCTTT  
67301 GAACTTGCTG TTTTCAATAT GGGCAGCACA AAGGTGAGAG ATACATATTA  
67351 ATAGTAGTAT GTATTACTCT TATACATTAG ATACCTATAT TTAAATGAAA  
67401 GGCCCAATTT GTAAACATAT ACATTCATAT TCTCTCTGC CCCAAGTTT  
67451 AGGAACATGT TAGGATATAG GAGACTTAAT TTATAATAAT GAGAGCATTT  
67501 TTTTATTTTA CTAAAGCCAT TTTTATAGTC AACTATCTTT TCTTATTTGT  
67551 GTGATTAGAA CTTAGAAAAA TATTTACTAG TTGAAGTTAT TATCAGTTTT  
67601 TAATTTAGTT CTTAAACTCA TTTCACTTCT AATAATTTCT GTTATAAATT  
67651 GCCAGCATT TAATGAAAA CTAAATGATGT AATAGGCATT TTCTTTATTT  
67701 GAACCTACCT CTTTTATTTT CTGAACCAA GAGAAAGATG GACTGGTGT  
67751 TGTGAAACAT TTTTAAAAAT GTAGTTTCAT TTATATTAGT TATGTTTGAT  
67801 AAATGTCTCA GTATTTTAT AATATGATAA GCCTGGGATT CTACTTTTAG  
67851 GGTTATTTGT ACTTTTGAGT AATATATAAA GTGACAATAT TAAGGTACAT  
67901 GATCAGCTCT TTCTATTTT ACTCGTAAAA ATTATGAAA TGAATAATTT  
67951 TGCTAACAAAC TTTGAAATTT CAACTTCTG GAAAATATGA AAATATTCAT  
68001 TGTTCATTAT GAATTTAAAT TGTAAGGTAT GAATGTGATT TGTCTGTACA  
68051 TCTTGTATCT TTTCCAAAA ATGATTCTGT ATCTTTTGA AAAAAAGCCGA  
68101 GAGTTGAAGA TAGTATATTT CTGGTAGTAC TGAATATTTA CTTACAGTTT  
68151 CTATCAAAAA TATATATTTG TTTCTAAAA TACTTGTTTT CCAGTTTTTA  
68201 TTTTTTTTAG AGAAAAATCT TAAGTCTCAG TTTCCCTAAT GAAAAAATA  
68251 AATTATAAAT AAAGCAAAAA TTGTATCCTA CAGCTTAGCT AGCTTAGATG  
68301 TTTGGCACCA GTTTGAATCA TGCTTTTATC AGCTGGCTCC ATGTAGTCTT  
68351 TCCAAAACATT TTGGCCTTTC CTGAGCAGCC CTTGTAGATA TTGTCTGTAT  
68401 GATGCATTTT GACACAAGGT GATATTTTT GTGATATCAA AATTCCACAT  
68451 TTACCCATTA GAGTTACAGC CCTGGGGTTC ACAGTACCAA GGGGGACCCA  
68501 GAGCTCAGG ATTGCCAGG CTCATTTTGC CGTGGAGTAT CAGTTTGTCT  
68551 TGAAATTTGT GGAAAAAATT CTAAGTTGAA TTCACTGGTA AGTAATTTTT  
68601 TAAAATTTCA TAATGCAGAT TACATCCAAA ATTTGATTTA AAAATTAATA  
68651 CATAAGACTG CAGAGAAATT CTGCATTCA ACTCCAATC TATCCAGACT  
68701 TCAGAAATAA CTTATCAGTT ATTTCTGTAA GCTTCTTGCT TACCTGGATA  
68751 CCTGACAGGT GAGATGGCTG TAGCAGACAC TGGCAGTTCC CTGCCCACAC  
68801 ACCTGTCCCT GTCCACAGCT GCACAAGGCA GCTCTGTGTG CAATTGCCAG  
68851 CATCTGCTCC TCTGTTCTCA GGAATCTTT GTTAGAAAA TGCTGCCATA  
68901 TTTGTTTCTC ACCTATTAGT CTTGTCTCCC AGTCAAGAGA ATAAATTTAT  
68951 GCAAGCAGAG ATTGTACTTT ACAGTATTTT GTCTTTGAGC TTGGCATTAG  
69001 GTTGCAATTG TAAAAATGTG GCATGGCTTC CTCATCCCCC AATAGGAACT  
69051 TTGCCAGCCC TTTTGTCTC ATGGAACCTC CTTTTTTGAA AAGAGCACCA  
69101 AAGGAGTAAA AATACTGTGG AGGGAGCAAC CCTCCTTTGC CATATGCTCT  
69151 CATTGGGAGA CATGTGGAGC AGTCTGAAGT CATTAGGCC ACTCTCTGGG  
69201 AGAGCACATC CTATGATGTT CTCCCAGCCT AGCCCCTTCC ACTGTGCTCA  
69251 AGTCCAAGCT GACCAGCTTT CTGACCACAG TGTAACAAA GATGATTGTC

FIGURE 3, page 22 of 42

69301 AGTGGGCCCC AGAATCCTAT ACCCAGA

**FEATURES:**

Start: 2132  
Exon: 2132-2314  
Intron: 2315-17055  
Exon: 17056-17182  
Intron: 17183-20983  
Exon: 20984-21071  
Intron: 21072-41719  
Exon: 41720-41831  
Intron: 41832-45391  
Exon: 45392-45550  
Intron: 45551-47878  
Exon: 47879-48031  
Intron: 48032-54612  
Exon: 54613-54720  
Intron: 54721-59290  
Exon: 59291-59458  
Intron: 59459-63791  
Exon: 63792-63942  
Intron: 63943-66164  
Exon: 66165-66346  
Stop: 66347

**CHROMOSOME MAP POSITION:**

Chromosome 1

**ALLELIC VARIANTS (SNPs):**

DNA

Position	Major	Minor	Domain
1722	G	C A	Beyond ORF(5')
1767	C	G A	Beyond ORF(5')
1840	C	G	Beyond ORF(5')
1857	T	G	Beyond ORF(5')
1945	G	T	Beyond ORF(5')
2007	A	C	Beyond ORF(5')
2769	C	G	Intron
3664	C	T	Intron
3827	G	A	Intron
4113	C	T	Intron
4337	A	G	Intron
4473	G	A	Intron
6455	T	G	Intron
6533	T	G A	Intron
6919	G	C	Intron
7305	G	A	Intron
7340	A	G	Intron
7466	A	G	Intron
7589	G	C	Intron
7810	A	C	Intron
9104	G	A	Intron
9503	A	T	Intron
9898	G	C	Intron
10196	T	C	Intron
12327	C	G A	Intron
13749	G	A	Intron
14150	T	C	Intron
14529	G	A	Intron
14653	G	A	Intron
15871	A	G	Intron
19244	G	A	Intron

19387	T	G	Intron
19447	C	G	Intron
20076	T	C	Intron
20492	T	-	Intron
20868	T	C	Intron
20941	T	C	Intron
21116	C	T	Intron
21701	G	A	Intron
21710	A	-	Intron
21826	C	T	Intron
21840	-	T	Intron
21841	-	C T	Intron
21843	-	C	Intron
22045	C	A T	Intron
22061	G	T	Intron
22348	-	A G	Intron
22682	A	G T	Intron
22783	-	T	Intron
23448	A	G	Intron
24960	G	A	Intron
24983	T	C	Intron
25390	T	C	Intron
26060	C	T	Intron
30245	C	G	Intron
33664	G	T	Intron
33883	C	A	Intron
34373	G	A	Intron
34558	G	T	Intron
43929	T	A	Intron
44309	T	- C	Intron
44997	T	G	Intron
46538	A	G	Intron
48153	T	C	Intron
48288	G	T	Intron
48412	G	A	Intron
48446	C	G	Intron
48456	G	C	Intron
48789	C	-	Intron
48859	G	C	Intron
49126	A	G	Intron
49378	T	G	Intron
49482	A	C	Intron
49741	G	A	Intron
49840	A	G	Intron
50102	G	A	Intron
50109	C	G T	Intron
50747	G	A	Intron
51272	G	A	Intron
52842	G	A	Intron
61837	A	G	Intron
62018	A	G	Intron
65562	A	G	Intron
65780	G	A	Intron
66092	G	A	Intron
66617	C	T	Beyond ORF (3')
66892	G	A	Beyond ORF (3')
67263	G	A	Beyond ORF (3')
67651	G	T	Beyond ORF (3')
67935	C	T	Beyond ORF (3')
69000	T	G	Beyond ORF (3')
69134	C	T	Beyond ORF (3')

FIGURE 3, page 24 of 42

Context:

DNA

Position

1722

TTGCCCACGCAGATGGCTGTTGATCTTTTCTGCAACAAATCCAGGAGTTTCTCCTTTTGT  
TTTTATAATTGCTCCAATAGATGCTTTAGGATTTAACTCTCTGCTTTTAAAGCAGAATC  
GCCATCCCAGGTGTGCAACCACGAAAAAATTAGACATCCGTGAGAGACAATGCCCTCCAT  
GGCCAGTTTCCAGGCAGAGAGAAGCAGCTCTGGGCTGACCGCCAAGGCTCCGGCCCGAG  
AGGGTCTTTAAGTGGAGTAACCAGTCTTCAAGACCCCGCTCCCAAGCCACCGACGCGTG  
[G, C, A]  
CGCTGCAGCCCTGGACCTGCTGGGGGCTCTTCTCGGACCCGCATGCTGACAGCGGGAC  
TGCCAATGGGCAGAGGTCGACCCCGGGTCCGCACAGCACCTCCGAGACCCAGTCCCA  
GCTCCCTCACTTCCGGCTCTCTGAGGCGGGCCCGCCAGTGCCGCGAGGCCAGCGCGG  
CGAGCTCTCCCGCAGCAGCGGCGGGACGGCCACACCCTGCGCGCCGCGCGGGCTCGGGT  
GGGTCTCCGCTCTGCGCCCTGCGCGCCGACGCCGACCCCGACGGCGCCCCAAACGCT

1767

AGTTTCTCCTTTTTGTTTATAATTGCTCCAATAGATGCTTTAGGATTTAACTCTCTGCT  
TTTTAAAGCAGAATCGCCATCCCAGGTGTGCAACCACGAAAAAATTAGACATCCGTGAGA  
GACAATGCCCTCCATGGCCCAAGTTTCCAGGCAGAGAGAAGCAGCTCTGGGCTGACCGCA  
AGGCTCCGGCCCGAGAGGGTCTTTAAGTGGAGTAACCAGTCTTCAAGACCCCGCTCCCAA  
GCCACCGACGCGTGACGCTGCAGCCCTGGACCTGCTGGGGGCTCTTCTCGGACCCCG  
[C, G, A]  
TGCTGACAGCGGGACTGGCAACTGGGCAGAGGTCGACCCCGGGTCCGCACAGCACCTCCC  
GAGACCCAGCTCCCAGCTCCCTCACTTCCGGCTCTCTGGAGGCGGGCCCGCCAGTGCCG  
CCGAGGCCAGCGCGGCGAGCTCCTCCCCAGCAGCGCGGGACGGCCACACCCTGCGCGCC  
GCGCGGGCTCGGGTGGGGTCTCCGCTCCTGCGCCCTGCGCGCCGAGCCGACCCCGAC  
GGCGCCCCAAACGCTGTTGCGCCGCGCGCCCGCCAGCCCGGCTCGCGCTGGTCCCGG

1840

TCGCCATCCCAGGTGTGCAACCACGAAAAAATTAGACATCCGTGAGAGACAATGCCCTCC  
ATGGCCCAAGTTTCCAGGCAGAGAGAAGCAGCTCTGGGCTGACCGCCAAGGCTCCGGCCCG  
AGAGGGTCTTTAAGTGGAGTAACCAGTCTTCAAGACCCCGCTCCCAAGCCACCGACGCGC  
TGACGCTGCAGCCCTGGACCTGCTGGGGGCTCTTCTCGGACCCGATGCTGACAGCGG  
GACTGGCAACTGGGCAGAGGTCGACCCCGGGTCCGCACAGCACCTCCCGAGACCCAGCTC  
[C, G]  
CAGCTCCCTCACTTCCGGCTCTCTGGAGGCGGGCCCGCCAGTGCCGCGGAGGCCAGCGC  
GGCGAGCTCCTCCCGCAGCAGCGCGGGACGGCCACACCCTGCGCGCCGCGCGGGCTCGGG  
TGGGGTCTCCGCTCCTGCGCCCTGCGCGCCGACGCCGACCCCGACGGCGCCCCAAACG  
CTGTTGCGCCGCGCGCCCGCCAGCCCGGCTCGCGCTGGTCCCGGTCTCGCCCCGAG  
CCCTCGATCTCCCGTGACTTCTCGGCCAGGCCGCTGCGCCTCTGGGACCATGTTGCGC

1857

CAACCACGAAAAAATTAGACATCCGTGAGAGACAATGCCCTCCATGGCCCAAGTTTCCAGG  
CAGAGAGAAGCAGCTCTGGGCTGACCGCCAAGGCTCCGGCCGAGAGGGTCTTTAAGTGG  
AGTAACCAAGTCTTCAAGACCCCGCTCCCAAGCCACCGACGCGTGACGCTGCAGCCCTGG  
ACCTGCTGGGGGCTCTTCTCGGACCCGCATGCTGACAGCGGGACTGGCAACTGGGCGAG  
AGGTCGACCCCGGGTCCGCACAGCACCTCCCGAGACCCAGCTCCAGCTCCCTCACTTCC  
[T, G]  
GCTCTCTGGAGGCGGGCCCGGCCAGTGCCGCGGAGGCCAGCGGGCGAGCTCCTCCCCAG  
CAGCGGCGGGACGGCCACACCCTGCGCGCCGCGCGGGCTCGGGTGGGGTCTCCGCTCCTG  
CGCCTGCGCGCCGACGCCGACCCCGACGGCGCCCCAAACGCTGTTGCGCCGCGCGCC  
CCGCCAGCCCGGCTCGCGCTGGTCCCGGTCTCGCCCCGACGCCCTCGATCTCCCGTGA  
CTTCTCGGCCAGGCCGCTGCGCCTCTGGGACCATGTTGCGCTGGCTGCGGGACTTCGT

1945

CAAGGCTCCGGCCCGAGAGGGTCTTTAAGTGGAGTAACCAGTCTTCAAGACCCCGCTCCC  
AAGCCACCGACGCGTGACGCTGCAGCCCTGGACCTGCTGGGGGCTCTTCTCGGACCC  
GCATGCTGACAGCGGGACTGGCAACTGGGCAGAGTGCAGCCCGGGTCCGCACAGCACCT  
CCCGAGACCCAGTCCCAGTCCCTCACTTCCGGTCTCTGGAGGCGGGCCCGGCCAGTG  
CCGCCGAGGCCAGCGCGGAGCTCCTCCCCAGCAGCGGCGGGACGGCCACACCCTGCGC  
[G, T]  
CCGCGCGGGCTCGGGTGGGGTCTCCGCTCCTGCGCCCTGCGCGCCGAGCCGACCCCG  
ACGGCGCCCCAAACGCTGTTGCGCCGCGCGCCCGCCAGCCCGGCTCGCGCTGGTCCC  
GGTCTCGCCCCGACGCCCTCGATCTCCCGTGACTTCTCGGCCAGGCCGCTGCGCCTCT  
GGGACCATGTTGCGCTGGCTGCGGGACTTCGTGCTGCCACCGCGGCTGCCAGGACGCG  
GAGCAGCCGACGCGCTACGAGACCCTCTTCCAGGCACTGGACCGCAATGGGGACGAGTG

2007 GCCACCGACGCGCTGACGCTGCAGCCCTGGACCTGCTGGGGGCCTCTTCTCGGACCCGC  
ATGCTGACAGCGGGACTGGCAACTGGGCAGAGGTCGACCCCGGGTCCGCACAGCACCTCC  
CGAGACCCAGCTCCCAGCTCCCTCACTTCCGGCTCTCTGGAGGCGGGCCCGCCAGTGCC  
GCCGAGGCCAGCGCGGCGAGCTCTCCCCAGCAGCGGCGGGACGGCCACACCTTGC GCGC  
CGCGCGGGCTCGGGTGGGGTCTCCGCTCTGCGCCCTGCGCGCCGACCCGACCCCGA  
[A, C]  
GGCGCCCCAAACGCTGTTGCGCCGCGCGCCCGCCAGCCCGGCCTGCGCTGGTCCCGG  
TCTCGCCCCGAGCCCTCGATCTCCCGTGA CTTCCTCGGCCAGGCCGCTGCGCCTCTGG  
GACCATGTTGCGCTGGCTGCGGGACTTCGTGCTGCCCACCGCGGCTGCCAGGACGCGGA  
GCAGCCGACGCGCTACGAGACCTCTTCCAGGCACTGGACCGCAATGGGGACGGAGTGGT  
GGACATCGGCGAGCTGCAGGAGGGCTCAGGAACCTGGGCATCCCTCTGGGCCAGGACGC

2769 TGGGGCCGCGACCGCGACCCCGGTAACAGAAGTGGGTCTATAATACGAAAGTCTACTGGT  
ATTTGTCCAGATAAAATGAGTGTGTGGACACTCTGGCCACGGGCACTGTAAATTTTT  
AAGACACTTTTGTCTGAATCCATCCCAGGTTCTTGT TTTCTGTTTAAATACCTTCAG  
ACATGTAATCCGTTT TAGCTGT CAGACTTCAGTGGGTCCCAAGTTTGTATAAAGCGCA  
ACATTCGATCTCTTTCGAAGCTGCTTTGTTACAGCAGCTATGTGTATTGTCTACTGTTT  
[C, G]  
AAAACGTGTTTGAAAACCAATCGCGTGT TTTCCCCACTTCCTGTTGAGAAGGAATGGCGGC  
ATTCCATTGTTTAAAGACATTCCTAGGTTAATGCCCTAGGTACATAAATTGATCTGAAGGG  
TTGACTTGACCTGCGACTGAGCAATTTCA TTTTCTCTGAGTCATCTTAACTGTGCCCTG  
AACTTCTGCCCCTTTAGTAGGGTGGAGATATGTGGAAC TTTCTCAACCCTGTTGAAGCGT  
TCCCTGACACTGGCATTCTCTTATCCAAAGAGGGAAAGTGATTAGGTTACTATGAGGGCC

3664 GCTGATTGTCCCAGAAATGGCCCAGTTGGAGTTCCCCACCATGTCCAATCATTGGCTGGA  
AGCAGCCCAGGAAAGGGACGACCTTGCTGCAGTGCATCAGCAGATGCCAGGGTTAGAGGC  
TAGAGAGTGGAAGTCAACTGTGTCTCTCACAGTAGGTGCCTTTGAAGGGAGATCTCAGTG  
GTACAACTCCATGGTCCCTACAATATACAAAAGCTCTTTGGAGTGTCAATGATTTTAA  
GATTGTAAAGGGATCCTGAGATCAAAAAGCTTGAGAATTGCTGCTGTATCACCATTTTAA  
[C, T]  
GTAAC TGCATCATATTTCTGTTATATGTTTGTGT CATAGTATATGTTACCAATTCTTTTAA  
AATCACCTTTTACTTTATTGATAGTTTAAAAACGATTGTAAGTGA AATTGCAATGGATGT  
CCTTTGTATTCA TTTTCTCATTCTGGTCCAGTTACTTTCTG TAGGATAAATTTTGAGGAGT  
GGACATTGCTGAGTCTGAAGGTAACACACATTTTAAACTGGGATACGTATTGCCTTTGCG  
AAACCTTAGACCCATTTTCACTCTTTTGACTGACAGTGCTTGCTTCTCCACATCCTCGCT

3827 GAAGGGAGATCTCAGTGGTACA ACTCCATGGTCCCTACAATATACAAAAGCTCTTTGGAG  
TGCTCAATGATTTTAAAGATTGTAAAGGGATCCTGAGATCAAAAAGCTTGAGAATTGCTG  
CTGTATCACCATTTTACGTA ACTGCATCATATTTCTGTTATATGTTTGTGT CATAGTATA  
TGTTACCAATTCTTTTAAATCACCTTTTACTTTATTGATAGTTTAAAAACGATTGTAAG  
TGAAATTGCAATGGATGTCCTTTGTATTCA TTTTCTCATTCTGGTCCAGTTACTTTGCTA  
[G, A]  
GATAAATTTTGAGGAGTGGACATTGCTGAGTCTGAAGGTAACACACATTTTAAACTGGGA  
TACGTATTGCCTTTGCGAAACCTTAGACCCATTTTCACTCTTTTGACTGACAGTGCTTGC  
TTCTCCACATCCTCGCTCATTTCAGGGTATCAGTCTTTGTAAAGTCTCCTATTCTGCAAGT  
GAAATTCCTTTTCA TTTCTGTCTTAGTCCATTTAGTGTTGCTATAGTGGAATATCTGAG  
ACAGGGTAATTTATAAAGAAAAGACATTTATTTAGCTCACAGTTCGCGAGGCTGGGAAGT

4113 CAGTTACTTTCTG TAGGATAAATTTTGAGGAGTGGACATTGCTGAGTCTGAAGGTAACACA  
CATTTTAAACTGGGATACGTATTGCCTTTGCGAAACCTTAGACCCATTTTCACTCTTTTG  
ACTGACAGTGCCTTGCTTCTCCACATCCTCGCTCATT CAGGGTATCAGTCTTTGTAAAGTC  
TCCTATTCTG CAGGTGAAATTCCTTTTCA TTTCTGTCTTAGTCCATTTAGTGTGCTAT  
AGTGGAATATCTGAGACAGGGTAATTTATAAAGAAAAGACATTTATTTAGCTCACAGTTC  
[C, T]  
GCAGGCTGGGAAGTTTAAAGAAGCGTGGTGCTGGCATCTGCTGGACTCCTGGGGAGGGCTT  
TCCTGCTGTGT CACAACATGGTGGAAAGTCAAAGTGGAAAGTGGACATGTGTGAAGAAGCA  
AAATCCGAGGGGTGCTCGCTTTTATAGCAACCCAGCCTCGAGGGA ACTGATCCATTACT  
GAGGGAAC TAATTCAGTCTCATGAGAGAGAGA ACTCACTACTGCAAGAATGACACC  
AAGCCATT CATGAGGGATCTGCCTCCGTAACCTGACACCTCCTGCTAGGTCCCTCCTCC

4337 CATTTAGTGTTGCTATAGTGAATATCTGAGACAGGGTAATTTATAAAGAAAAGACATTT  
ATTTAGCTCACAGTTCGCGAGGCTGGGAAGTTTAAAGAAGCGTGGTGCTGGCATCTGCTGG

FIGURE 3, page 26 of 42



ACTCCTGGGGAGGGCTTTCCTGCTGTGTCACAACATGGTGGAAAGTCAAAGTGGAAAGTGG  
 ACATGTGTGAAGAAGCAAATCCGAGGGGTGTCTGGCTTTATAGCAACCCAGCCTCGAG  
 GGAAGTGTATCCATTACTGAGGGAACATAATTCAGTCTCATGAGAGAGAGAACTCACTCACT  
 [A, G]  
 CTGCAAGAATGACACCAAGCCATTTCATGAGGGATCTGCCTCCGTAACCCTGACACCTCCT  
 GCTAGGTCCCTCCTCCCAACACGGCCACATCAGGGATCAGACTTCAACATGAGTTTTTGT  
 GGGGACAAACAAACGTAGCACTTGCTTTGCCTTTTGGTTCTATTACATCCTCCACAGG  
 ATTGCATTATGCCTACCCATTTGGTGAGGGCAGTCTTCTTAATTGGTTTACTGATTCAA  
 ATGCTACCCTCCTCCAGAGACATCCTCACAGACACACCCAGAAATCATGTTTTACCAGTT  
 4473 TTCCTGCTGTGTCACAACATGGTGGAAAGTCAAAGTGGAAAGTGGACATGTGTGAAGAAGC  
 AAAATCCGAGGGGTGTCTGGCTTTATAGCAACCCAGCCTCGAGGGAAGTGTATCCATTAC  
 TGAGGGAACATAATTCAGTCTCATGAGAGAGAGAACTCACTCACTACTGCAAGAATGACAC  
 CAAGCCATTTCATGAGGGATCTGCCTCCGTAACCCTGACACCTCCTGCTAGGTCCCTCCTC  
 CCAACACGGCCACATCAGGGATCAGACTTCAACATGAGTTTTTGTGGGGACAAACAAAC  
 [G, A]  
 TAGCACTTGCTTTGCCTTTTGGTTCTATTACATCCTCCACAGGATTGCATTATGCCTAC  
 CCATTGGTGAGGGCAGTCTTCTTAATTGGTTTACTGATTCAAATGCTACCCCTCCTCCA  
 GAGACATCCTCACAGACACACCCAGAAATCATGTTTTACCAGTTATCTGGGCATCCCTTA  
 GTCCAGACGAGTTGATACATAAAATTAACCATCACACATGGGATAGAATTAGGATTACAC  
 AGTCAACCTTTATGGGAGAAAATTCAGAGGCATGTCAGGGGTTTATGTAATGTCAAGGA  
 6455 TGTTTATTGCATTGAGTGGAAATCAGGATTTCACTCCATTAAGTAATTCCTCTGTTAACAA  
 AGAGGGTTCATTTTCAATTTTATTTTCAATATATGCTTTTTTTTTTTTTTTTCTGGAGAC  
 AGAATCTTGCTCTATACCAAGGCTGGAGTGCAGTGGTGCAGTCTCGGCTCACTGCAGCC  
 TCTGCTTCCCTGGATTCAAGCGATTCTTGTGCCTCAGCCTCCCAAGCAGCTGAGATTACAG  
 GCACATGCCACCACACCTGGTTAACTTTTGTATTTTCTAGTAGAGATGGGATTTTGCCTAT  
 [T, G]  
 TTGGTCAGGCTGGTCTTGAATTCCTGGCCTCTAGTGATCTGCCTGCCTCTGCCTCTGAAA  
 GTGCTAAGATTACAGGCATGAGCTACCATGGCCAGCCCATTTCTTAAATATTTTAAATGT  
 CAGACATGTTATGGTTTCTGGCACAATATTAAGAAGACATGATATGAAATCACAGGGTGA  
 ATTTTAGGGCATCACAACAGAAAGATTATGGTATAAGAAAAACAATGGAATTCCAACTAC  
 ATTTCTGTCAAATGTTCTAAAATATATAAAATCTGTATCTTTTGTGTTCTCTCCTGATTT  
 6533 TTATTTCAATTAATATGCTTTTTTTTTTTTTTTTTTCTGGAGACAGAATCTTGCTCTATCAC  
 CAAGGCTGGAGTGCAGTGGTGCAGTCTCGGCTCACTGCAGCCTCTGCTTCCTGGATTCAA  
 GCGATTCTTGTGCCTCAGCCTCCCAAGCAGCTGAGATTACAGGCACATGCCACCACACCT  
 GGTAACTTTTGTATTTTCTAGTAGAGATGGGATTTTGCCATGTTGGTCAGGCTGGTCTT  
 GAATTCCTGGCCTCTAGTGATCTGCCTGCCTCTGCCTCTGAAAGTGCTAAGATTACAGGC  
 [T, G, A]  
 TGAGCTACCATGGCCAGCCCATTTCTTAAATATTTTAAATGTGACACATGTTATGGTTTC  
 TGGCACAATATTAAGAAGACATGATATGAAATCACAGGGTGAATTTAGGGCATCACAAC  
 AGAAAGATTATGGTATAAGAAAAACAATGGAATTCCAACTACATTTCTGTCAAATGTTCT  
 AAAATATATAAAATCTGTATCTTTTGTGTTCTCTCCTGATTTATATTCTAAATTTGATGT  
 TATCCTTCTCTGCAGAAATAAAGTGTCTGAAAGAATGAAAAAATGGAAGAATTCCTTAG  
 6919 ATGAAATCACAGGGTGAATTTTAGGGCATCACAACAGAAAGATTATGGTATAAGAAAAAC  
 AATGGAATTCCAACTACATTTCTGTCAAATGTTCTAAAATATATAAAATCTGTATCTTTT  
 GTGTCTCTCCTGATTTATATTCTAAATTTGATGTTATCCTTCTCTGCAGAAATAAAGTG  
 TCTGAAAGAATGAAAAAATGGAAGAATTTCTTAGTAAGGTATAAAATACCCCTTTCTATC  
 TTTGTAGCATTCTAAGCCTTTTGTACCTTTCCAACTCCCAACATGCCATATCCCTGA  
 [G, C]  
 TAGGCCACAGCCATGTACATTGATCCCTTTATTTTCTTCTCTGCCTGAGATTTCTCTC  
 ATTCCCCCTTCTCTGCCTGGTATATGATTGCCCATTTGTTAAGGCCCAACTCACCTTTA  
 TAATCTTCTAGCCCACTTTCTTTATCGGTATTCAGAAAAACAAAAGAAGCTTCCACA  
 AGACAACATTCTGTAATACACTGCTTAACCTCTTTTGACCCTGCTGAGTTCAAAAATCTT  
 ATCTTTTAAAGGATTGAATGGAGTCCACCAAGGTATCTATATTTGACAGGATTTATGAAA  
 7305 GATTGCCCATTTGTTAAGGCCCAACTCACCTTTATAATCTTCTAGCCCACTTTCTTTA  
 TCGGTATTCCAGAAAAACAAAAGAAGCTTCCACAAGACAACATTCTGTAATACACTGCT  
 TAACTTCTTTTGACCCTGCTGAGTTCAAAAATCTTATCTTTTAAAGGATTGAATGGAGTC  
 CACCAAGGTATCTATATTTGACAGGATTTATGAAAACAAAAGGATTTGTTGAGAAAGTTT  
 GAAGCCTAACTCTGAAACGTGGATCATAGTGTTTACTACACATTAACCTGTTTTAGTGGAT

FIGURE 3, page 27 of 42

[G, A]  
TAATAGTTATTATTATAGGCTGTGGAATCAGAACAGGGTTCAAATGTTTTACCGCTTGC  
TAGACTGTGGCCTTGGGCATGTTATTAAATGCCTGGAGGCCTCAAATGTTAACTAGGAAT  
GGTAAGACCTACCCAGTAACCTAGCATAAATAGTAAATTCATTCAATTAATGTTTTCAA  
CAGTGCCAGACATTGTTAATGAAGTGGGATATAGTGGTGAACAACACTGACAGCGTTC  
TTCATTGTATTCTCAAACCCCTCCCTATAGTAAGTAGGTCTGTGTGTGTGTGTAGGTGCA

7340 TAATCTTCCTAGCCCACTTTCTTTATCGGTATTCCAGAAAAACAAAAGAAGCTTCCACA  
AGACAACATTCTGTAATACACTGCTTAACCTCTTTTGACCCTGCTGAGTTCAAAAATCTT  
ATCTTTTTAAGGATTGAATGGAGTCCACCAAGGTATCTATATTTGACAGGATTTATGAAA  
ACAAAAGGATTTGTTGAGAAAGTTTGAAGCCTAACCTGAAACGTGGATCATAGTGTTTA  
CTACACATTAACCTGTTTGTAGTGGATGTAATAGTTATTATTATAGGCTGTGGAATCAGAAC  
[A, G]  
GGGTCAAATGTTTTACCGCTTGCTAGACTGTGGCCTTGGGCATGTTATTAAATGCCTG  
GAGGCCTCAAATGTTAACTAGGAATGGTAAGACCTACCCAGTAACCTAGCATAAATAGTA  
AATTCATTCAATTAATGTTTTCAAACAGTGCCAGACATTGTTTAAATGAAGTGGGGATATA  
GTGGTGAACAACACTGACAGCGTTCTTCATTGTATTCTCAAACCCCTCCCTATAGTAAGT  
AGGTCTGTGTGTGTGTGTAGGTGCATGGGAATAAAAAATAAAGCAAATAATGAACAG

7466 TTAAGGATTGAATGGAGTCCACCAAGGTATCTATATTTGACAGGATTTATGAAAACAAA  
GGATTTGTTGAGAAAGTTTGAAGCCTAACCTGAAACGTGGATCATAGTGTTTACTACAC  
ATTAAGTGTGTTTGTAGTGGATGTAATAGTTATTATTATAGGCTGTGGAATCAGAACAGGGTT  
CAAATGTTTTACCGCTTGCTAGACTGTGGCCTTGGGCATGTTATTAAATGCCTGGAGGC  
CTCAAATGTTAACTAGGAATGGTAAGACCTACCCAGTAACCTAGCATAAATAGTAAATTC  
[A, G]  
TTCATTTAATGTTTTCAAACAGTGCCAGACATTGTTTAAATGAAGTGGGGATATAGTGGTG  
AACAACTGACAGCGTTCTTCATTGTATTCTCAAACCCCTCCCTATAGTAAGTAGGTCT  
GTGTGTGTGTGTAGGTGCATGGGAATAAAAAATAAAGCAAATAATGAACAGGGTAAT  
TTCAAAAAGCAGAAAGAGCTATTCAACAAACTACCTGCCTTTTATTAGATGAACTCTC  
AACTCTATGGTTTGTCTCTCCTGTCAATTCTGTTAAATGCTGTGAGCCTGTTTTCTCTTA

7589 AACTGTTTTAGTGGATGTAATAGTTATTATTATAGGCTGTGGAATCAGAACAGGGTTCAA  
ATGTTTTACCGCTTGCTAGACTGTGGCCTTGGGCATGTTATTAAATGCCTGGAGGCCTC  
AAATGTTAACTAGGAATGGTAAGACCTACCCAGTAACCTAGCATAAATAGTAAATTCATT  
CATTTAATGTTTTCAAACAGTGCCAGACATTGTTTAAATGAAGTGGGGATATAGTGGTGAA  
CAACACTGACAGCGTTCTTCATTGTATTCTCAAACCCCTCCCTATAGTAAGTAGGTCTGT  
[G, C]  
TGTGTGTGTAGGTGCATGGGAATAAAAAATAAAGCAAATAATGAACAGGGTAATTTTC  
AAAAAGCAGAAAGAGCTATTCAACAAACTACCTGCCTTTTATTAGATGAAACTCTCAAC  
TCTATGGTTTGTCTCTCCTGTCAATTCTGTTAAATGCTGTGAGCCTGTTTTCTCTTATCA  
CCCTGGCCACGACTTCTGTCTTTCTGCTTGGTCTGTAGACTCTAACCAAGGCTCATT  
CTCTGCCTGGCTATCTGCCTTCTGTGGCTCTTGCCACTACCTACATTTTCTGTGTGTGCA

7810 CTGGGGATATAGTGGTGAACAACACTGACAGCGTTCTTCATTGTATTCTCAAACCCCTCC  
CTATAGTAAGTAGGTCTGTGTGTGTGTGTAGGTGCATGGGAATAAAAAATAAAGCAA  
ATAATGAACAGGGTAATTTCAAAAAGCAGAAAGAGCTATTCAACAAACTACCTGCCTTT  
TATTAGATGAAACTCTCAACTCTATGGTTTGTCTCTCCTGTCAATTCTGTTAAATGCTG  
TCAGCCTGTTTTCTTATCACCTGGCCACGACTTCTGTCTTTCTGCTTGGTCTGTAG  
[A, C]  
CTCTAACCCAAGGCTCATTCTCTGCCTGGCTATCTGCCTTCTGTGGCTCTTTGCCACTAC  
CTACATTTTCTGTGTTGCACAGGAAGGACCATTCCTGTGGACCATAAAATTCCTTTT  
TGAAAGAATTCATTCTTGATTGGGCCACAGCACATCTGTGAAACAGCATTAGACATTTG  
CCACTGCTCAGCAGCTCTGGGGGAAATGTTTACTGAGAAGCGTACAGTAGTTTTTTTGA  
CTAACCATGGTGCAACCTCCTCCCAGAGGGAACCTATGAGTATTTCAAGGACATGTGAT

9104 TTAAACGAATTATTGTAGAAACAGAAAAACAAATACTGTGTTCTCATTACAGGGGGAGC  
TAAACCTTGGGTAATGGGGCATAAAGATGGGAACAATAGACACTAGGGGACTCCAAAAGG  
GGGAGGGAGGGAGGAGGGCAAGGGCTGGAAAGCTTCTACTGGGTACTTTGTTCAACAC  
CTGGGTGATGGCAGATTAGGAGCTCAACCCAGTATCACACAGTATACCCTTGTAACA  
AGCTGATGGTGTAAACCCCTGAATCTACAATAAAATTATTTTATTTTAAAAAATCATTATA  
[G, A]  
GGATTTTTTAAAAAGAAGGATTCTAGACAGGTGCAGCCAAACAATTTTTTTTAAATGTTG  
GCAGGCCGCCACCGCAGTCATTATGCTGCAATAGCCCATGTCCCAACATTCCCAACCT

ACTTCTCTCCAAAAGAGAAGCTATACTTTTCAGATGGCCCTGTGCTGGGTTCTCCCTGGAA  
GTTTCTGGGGAAAGGGGCTTGAGTTGCCCCGACTGGACTCTTCTGGAGTGGGAGCCGGG  
GCTTCTGATCAGACGTGAGTGAGGCAGGAACCTCCGCGGTCTCCAGCGCAGCCAGAGTG

9503 CATGTCCCAACATTCCCAACCTACTTCTCTCCAAAAGAGAAGCTATACTTTTCAGATGGCC  
CTGTGCTGGGTTCTCCCTGGAAGTTTCTGGGGAAAGGGGCTTGAGTTGCCCCGACTGGAC  
TCTTCTCTGGAGTGGGAGCCGGGGCTTCTGATCAGACGTGAGTGAGGCAGGAACCTCCGCGG  
TCTCCAGCGCAGCCAGAGTGCGGTCCACGCAGGTCCCGGTCTGCGCGCTCGCGCC  
TTTGCCTGAAGCCGTTAGGATGAGCCCTCTCCTTCCAGAGCTTTAACCGATGAAGGTGC  
[A, T]  
TTGTGTTTGGCGCCCCCTGAGGAGGATGCTGTCTTAGGCCCTTCCCACTGGACGTGTGTG  
GTGGGCAGAGATCCCGTTCGTGCGTGCCTTCCACCCGCTGGGGCTCACTCAGGCCGC  
GGAGCTGCCAGGGGAGACATCTCGATGGACTCCCTCTACGGAGATCTCTTTTGGTACCTG  
GACTATAACAAGGATGGGACCTTGGACATTTTGGAGCTTTCAGGAAGGCCTGGAGGATGTA  
GGGGCCATTCAATCTCTAGAGGAAGCGAAGGTGGGTCTCACTGGGGCTGTAATCAGAGAG

9898 ACCCCGCTGGGGCTCACTCAGGCCGCGGAGCTGCGAGGGAGACATCCTCGATGGACTCCC  
TCTACGGAGATCTCTTTTGGTACCTGGACTATAACAAGGATGGGACCTTGGACATTTTGG  
AGCTTCAGGAAGGCCTGGAGGATGTAGGGGCCATTCAATCTCTAGAGGAAGCGAAGGTGG  
GTCTCACTGGGGCTGTAATCAGAGAGACGTTGGGGCTGGGAGCCCTGGAGAGGCATTGGG  
CAGAGAGGGCAAAATTTACATGTTGTCAAGCTTGACCTGGGCCCCACTGCAGTGTTCAAGT  
[G, C]  
GTTGACCAGCGTTACCGTTTATTAAGAATAACAACACAGCTAACACATTTCTCAAGTATT  
TTTCTCCGTTTCTCCTTGGCTGTAGTAAAATCTCCAACCTTCAGATTGCTCTCAAGATGT  
TGGCTACATACAGCCTTGTCTTAGGAGTCACCTTGTCAATGTGCTCACCTGTCTATTAGT  
CAGCCAGAGGGGCGTCTAGGCTAAAGATGCGCCCTCCCCAGTTCAGAGAACTGGAATAAT  
CACTCTACGTGTATTTGGGAGTGGGGTGGTGATTGGAAATTTTCTGATGTTATGTTTGG

10196 GTGGTTGACCAGCGTTACCGTTTATTAAGAATAACAACACAGCTAACACATTTCTCAAGT  
ATTTTCTCCGTTTCTCCTTGGCTGTAGTAAAATCTCCAACCTTCAGATTGCTCTCAAGA  
TGTTGGCTACATACAGCCTTGTCTTAGGAGTCACCTTGTTCATGTGCTCACCTGTCTATT  
AGTCACCCAGAGGGGCGTCTAGGCTAAAGATGCGCCCTCCCCAGTTCAGAGAACTGGAAT  
AATCACTCTACGTGTATTTGGGAGTGGGGTGGTGATTGGAAATTTTCTGATGTTATGTTT  
[T, C]  
GGTTTCTGTTCTGGAAGGGGGCAGTGGAAGTGGCTTTTACTCTCGGGTTTCACTAGTGC  
TGAGGTTTCTCATAATATGCCTTAATTGATAGACCTAGTTATCAGTACCGAGCTTAGG  
CTAACCCCTTCTCTCCCCAGAAGGCTAACCTACAGGCTCCTTCTCAGCATGTTGTGCTTC  
GTACATACTCCTATTGCAGTATTTCCAAGTCATTTTTCATTTGGAATTTATTATTGTATA  
TAATAATTACTTTATAAGTATATTTTGCTCTTTGGATGTTTGACCCGGTAGACTGGGAGAT

12327 GTCATGTTATTTAATGCCTGGAGGCCCTCAAATGTTAACTAGGTAATGGTAAGACCTACCC  
AGTAACTTAGCATAAATAGTAAATTCATTCAATTTAATGTTTTCAAACAGTGCCAGACATT  
GTTTAATGAAGTGGGGATATAGTGGTGAACAACACTGACAGCGTTCTTCATTGTATCTC  
AAAACCCCTCCCTATAGTAAGTAGGTCTGTGTGTGTGTGTAGGTGCATGGGGAATAAAAAA  
TAATAAGCAAATAATGAACAATAAAATATTTTATTTAAAAAAAAGAAATGATACTTAC  
[C, G, A]  
TTGTGCTGTTAAGATACAAAAGCAATAACTTTTTATTGTGAAAATAGTCTGTTTTTGAAC  
AATATATTGTTTTGTTTTTCTGTGAAAGTTGAGAACTAAATATACGAAGAGATAATG  
GTCAGACCATAAAATAAAATAGAACTTTGACTCAAAATTTACAGCAGTCTGCCAGAAAA  
CCAGCCCTTTATCTAAAAATAAACAGACCAGGAACAGCCTGTTATGTCAGACTTATAGG  
AAGTCAGGTTGCTATCTCTAGAGACAATACACAAAGCTATGCAATAACTGCTGTAACAGC

13749 TACAGGCGTGAGCCACCATGCGCCGAGCCATAGACTATATATTTTGTATCTGATAACTGG  
TTCAGCTACTAAGTGACTAACAGGCAAGTAGCATCTATAGTGTGGATATGCTGGACAAAA  
GGACATTACCTCCTGGGCAGGATGGCACAGAATGTTGAGAGATTTTATCATGCTACTCA  
GAATGGTGTGCAATTTAAACTTATGAGTTGTTTGTCTTCTGGAGTTTTCATTTAATAGT  
TCAGACCATTGATTGACCGCAGGTAAGTGAAGTGTGGAGAGTGAAGTGTGGATAAGGG  
[G, A]  
GGACTATTGTATTGTTAAGTCAGACTCATTAGGCAATCATAACTCTTGATTGTCATCAG  
AAATGCTGCAGAAATATGGGTTAAAAAAAAGTGTCAAAATAGGGTCAGGGATGTCCTT  
TAACTTGTACTTCCAAAATGTTAGTGAAGTGTGGCCCCAAAGAGTGAAGGAACAAA  
TGACTAAGAGAAAAATCTGTTTTTCAAGATGACAGATTAAAAAAGAAGCAACTTGTGAAA  
CACTGAAAATCTCTCACTTGTAAAGATAACACAAAAGTGGCTAAAGTGGTTGGAATGAA

14150 ATAGGGTCAGGGATGTCCTTTAACTTGTTACTTCCAAAATGTTAGTGAAAACGTGGCCC  
CAAAGAGTGAAAGGAACAAATGACTAAGAGAAAATCTGTTTTTCAGGATGACAGATTAAA  
AAAGAAGCAACTTGCTGAAACACTGAAAATCTCTCCACTTGTAAAGATAACACAAAACCTGG  
CTAAAACCTGGTTGGAATGAATATGGCCAACCTCAAGTCTGCACAGAACTAACTTGGTGATG  
TTACAGCCCCAAATTTCCACCACATATTTTATACTAACTCCCCCGGATTTTCACACATGA  
[T, C]  
CTGTGAGGTAGCATGAAGAGGTAACCTATGCATGCCTAAGGACTTGGGAGACCTCCCCATT  
TCCTTCCACCAATCACCCACTAATCCAGAATCCGCCCCCAAACCTTTTCTAATAACTAC  
CTTAAAGCCAGCATAGGGGAGACAGATTTGAGCTGGACTCCTGTCTTCTGTGGGTACCT  
TGCAATAAAAAGCTTTTCTTTTCTCAACACCTGGTATTATAGTATTGACTTCTAGTTCAT  
CGGGCAGCAAGCCCCCTTTGGTGGTGACTATTCTGTTCGCTGATATTTCCATTGGCCA

14529 ACTAATCCCAGAATCCGCCCCCAAACCTTTTCTAATAACTACCTTAAAGCCAGCATAGGG  
AGACAGATTTGAGCTGGACTCCTGTCTTCTGTGGGTACCTTGCAATAAAAAGCTTTTC  
TTTTCTCAACACCTGGTATTATAGTATTGACTTCTAGTTCATCGGGCAGCAAGCCCCCTT  
TGGTCGGTGACTATTCTTGTTCGCTGATATTTCCATTGGCCAAAATATAAACCTCTTAGA  
TGAAACTTCAGTACGTAAATGGCGCCACAGAATGCTGTGACATTTTCTCTTGGATTATA  
[G, A]  
CAGGTTACTTTTACTGAATACCGTAGGCAGTTATAACACACTAAGTATTTGTGTATCTAAA  
CATAGAAAAGATACAGTAAAATATGGTAATTTTTTTCAACTTTTAGTTGAGATTTGGAG  
GGTATGTGCACATTTGTTACAAGGTATATTGCATGATGCTGAGGTTTGGGGTACAATTG  
AACCTGTACCCAGGTAGTGAGCATAGTACCCAATCGATAATTTTCAACCTTGTCCA  
TTCCCTCCCCGTTCTTGTAGTCCCCAGTTTCTGCTTTTCCCATCTTTATATCCGTGTGCA

14653 CTCAACACCTGGTATTATAGTATTGACTTCTAGTTCATCGGGCAGCAAGCCCCCTTTGGT  
CGGTGACTATTCTTGTTCGCTGATATTTCCATTGGCCAAAATATAAACCTCTTAGATGAA  
ACTTCAGTACGTAAATGGCGCCACAGAATGCTGTGACATTTTCTCTTGGATTATAGCAG  
GTTACTTTTACTGAATACCGTAGGCAGTTATAACACACTAAGTATTTGTGTATCTAAACAT  
AGAAAAGATACAGTAAAATATGGTAATTTTTTTCAACTTTTAGTTGAGATTTGGAGGGT  
[G, A]  
TGTGCACATTTGTTACAAGGTATATTGCATGATGCTGAGGTTTGGGGTACAATTGAACC  
CTGTCAACCAGGTAGTGAGCATAGTACCCAATCGATAATTTTCAACCTTGTCCATTCC  
CTCCCCGTTCTTGTAGTCCCCAGTTTCTGCTTTTCCCATCTTTATATCCGTGTGACCCC  
ATGTTTTGCTCCCATGTGTATGTGAGAACTTGTGGTGTGGTTTTCTATTTCTGCGTTG  
ATTGCTTAGGATAATGGCCTTCAGCTGCATCCATGTTGCTGCAGAGGACGTGATTTAT

15871 AGGAGTTTATCAATTTTATTAGTCTTTTCAAAGAACCATCTTTGGCTTTGTTAATCCTC  
CCATGGTGTGTTTTCTTTCTCATTACTTTTTGCTCTTATTTCCTTCAACTTCTTTTT  
GCTTAATTTTAAATAATTTCTTGAGATTGAGATAAGCCTCAATGATGGGTCACCGATTT  
CCAGTCTTTCTTCTTTCTAATTTATGCATTTTAAACCAGAAATCTTCTCTAAGTGTAGC  
TTTAGTTGTCAGCTACAAGTTTCAGATCTGTCTCTCAGTCTGGAGGTTGGAGATCTGACC  
[A, G]  
TGACCATGAAACCATCCAGTCACAATGTGGCATTATTTTTTTAATTTTTTTTTTTTTTT  
TGAGATAGAGTTTCACTCTTATTGCCTAGGCTGGTGTGCAATGGTCCGATCTCGGCTCAC  
AGCAACCTCCACCTCCAGGTTCAAGCGATTCTTTGCTCAGCCTCCCAAGTAGCTGGG  
ATTACAGGCATGCGCCACCATGCCCACTAATTTTGTATTTTGTAGAGATGGGGGTTT  
TCCATGTTGGTCAGGTTGGTCTTGAACCTCCGACCTCAGGTGATCCGCCCACCTCAGCCT

19244 GTGGCATTATTGGTTCATATTTTTATTTTTTAGACTTCCTTAATGCAAAACATATACAGT  
TGATCCTCATTATTTGGGGATTCTGTATTTGCAAATTTGCCTACTCAATAAAATTTATCC  
CCAAAGTAACCCCAAAATATATACTCACAGTACTTTCCAGGCATTCATGGACATGCACA  
GAGCAGTGAAAAACTTGAGTTGCTCAGCATGTACATTCTAGCTAGTAGAATAAGGCAAT  
ACTCTGCCTTCTTGTTCAGCTCTCATACTATTAAGTACAGTATCCCTTTCAAGGTCT  
[G, A]  
TTTTGTGCCAGTTTTTGCATTTTTGTATTTTTGTTGGTAATTTCTTTTTTAAATGTTCC  
CCAAAGGTAGTGCTGAAGTGCTGTCTAGTGTTCTTAAGTGCAAGAAAGCCATAGCATGCC  
TTATGGAGAAAATATATGCGTTGGATAAGCTTTGCCCAAATTCAATGTTAGTGAATCAA  
CAGCACACATTAATGAGGTGCCCTTCAAACAGAAAACAGACATAAGACATGGTTATGTATT  
AATCAGTTGATGAAAGTGTTGTAATCAGAGGCTCACAGGAACCTAACCTGTTTTTCTGTC

19387 CTCACAGTACTTTCCAGGCATTATGGACATGCACAGAGCAGTGAAAACTTGAGTTGC  
TCAGCATGTACATTCTAGCTAGTAGAATAAGGCAATACTCTGCCTTCTTGTTCAGCTC

TCATACTATTAAGTAGCAAGTATCCCTTTCAAGGTCTATTTTGTGCCAGTTTTTGCATTT  
TTGTATTTTTTGTGGTAATTTCCCTTTTAAAAATGTTCCCAAAGGTAGTGCTGAAGTGCT  
GTCTAGTGTTCCCTAAGTGCAAGAAAGCCATAGCATGCCTTATGGAGAAAATATATGCGTT  
[T, G]  
GATAAGCTTTGCCCCAAATTCAATGTTAGTGAATCAACAGCACACATTAAATGAGGTGCC  
TTCAAACAGAAAACAGACATAAGACATGGTTATGTATTAATCAGTTGATGAAAGTGTGTA  
ATCAGAGGCTCACAGGAACCTAACCTGTTTTTCTGTAGGAACAATGGTTTGGTATTTG  
CTAATTGAGTGTTTGAATGAATATAGAAGTTTATGGAAGATGATTGCTGTGAATAATGA  
GAATTAACCATATCTCTTTAAGAGTGCATTTCTAAAGGAGAATATTGAGAAGGTATTTG

19447 TCAGCATGTACATTCCTAGCTAGTAGAATAAGGCAATACTCTGCCTTCTTGTTCAGCTC  
TCATACTATTAAGTAGCAAGTATCCCTTTCAAGGTCTATTTTGTGCCAGTTTTTGCATTT  
TTGTATTTTTTGTGGTAATTTCCCTTTTAAAAATGTTCCCAAAGGTAGTGCTGAAGTGCT  
GTCTAGTGTTCCCTAAGTGCAAGAAAGCCATAGCATGCCTTATGGAGAAAATATATGCGTT  
GGATAAGCTTTGCCCCAAATTCAATGTTAGTGAATCAACAGCACACATTAAATGAGGTGCC  
[C, G]  
TTCAAACAGAAAACAGACATAAGACATGGTTATGTATTAATCAGTTGATGAAAGTGTGTA  
ATCAGAGGCTCACAGGAACCTAACCTGTTTTTCTGTAGGAACAATGGTTTGGTATTTG  
CTAATTGAGTGTTTGAATGAATATAGAAGTTTATGGAAGATGATTGCTGTGAATAATGA  
GAATTAACCATATCTCTTTAAGAGTGCATTTCTAAAGGAGAATATTGAGAAGGTATTTG  
CATAATTTCTTTACTAACAGATGCTGCCTCTCACTGTCTTACATGGTCCAGATTCTCAT

20076 TCTCTCAGAATCCTGTCTCTCCTCCAGGGTCCTTTCTCCAAGAAAGTCTATCCTTTAC  
CACTAACAGTAATTTTGGTCTTCTCTTTTTCTGGAGAAAGTCAAGTGTATGCTGCTTC  
AGCACCAGACCTCTCTTACTTTGTTTTGTTTCATTCTTTTTCATGTACAGTAGTCTTAG  
GATTCTCATGAGCCTGTGAGCTGCTAGAAGGAAATACAGCAGTGCTTACATTTATGCTT  
CTATTTTATTTCTATTTTCTCTCTCTGCTTCTGATTGTTCTCCTTCTGTCCACAAACA  
[T, C]  
GCTCTAATTTCCCTAGTATTAAAAATTTTCTGTCTTTTGTGTTCTTTTATCCTTGCTCC  
CTTATTTTTTACTGCCAGATTTTATTTTTATTTATTTATTTTGTAGATGGAGTCTCACTC  
TGTCACCCAGGCTGGGGTGCAGTGGCGCGATCTCAGCTCACTGCAACCTCCGCCTCCACG  
CTTCAAGCAATTTTCTCTTTTAGCCTCCCAAGTAGCTGGGATTATGGGCACCTGCCACC  
ATGCCTGGCTGATTTTCTATTTTATAGTAGAGACGGGGTTTACCATGTTGGCCACACTG

20492 CACTCTGTCAACCCAGGCTGGGGTGCAGTGGCGCGATCTCAGCTCACTGCAACCTCCGCCT  
CCAGCTTCAAGCAATTTTCTCTTTTAGCCTCCCAAGTAGCTGGGATTATGGGCACCTG  
CCACCATGCCTGGCTGATTTTCTATTTTATAGTAGAGACGGGGTTTACCATGTTGGCCA  
CACTGCTCTCTAAGTGTGACCTCAGGTGAACACCCGCCTCAGCTCCAAAAGTGCTGG  
GATTGCAAGGTGTGAGTCACTGTGCCTGGCCTTTTACTGCCAGATTTTAAAAGAATAGTC  
[T, -]  
GTGCTTTAGCTCTATTTCTCTACTTCTCTTTAACTCAGTCATATATGATGTTT  
TGATAGTAAATGCTAGTAATTTATTAATAATGTAGAAATAGGTACTTTTAAAATGAAT  
AGATCCTACTTTAATTGAATTTATCTTGGAGTTAGAATATCTTGATTTGGATTTTGTTC  
TGCTACTTCTTAATTACATTACTTGGTAAGGCCACTTGTGAAGTCACTCTTTGGAGGA  
ATATTATTTATCTATAAGGCTGTTACAATTACTGAATTTTAAAAATGTGTATTTATTTT

20868 TAGTAATTTATTAATAATGTAGAAATAGGTACTTTTAAAATGAATAGATCCTACTTTAAT  
TGAATTTATCTTGGAGTTAGAATATCTTGATTTGGATTTTAGTTCTGCTACTTCTTAATT  
ACATTACTTGGTAAGGCCACTTGTGAAGTCACTCTTTTGGAGGAATATTATTTATCTAT  
AAGGCTGTTACAATTACTGAATTTTAAAAATGTGTATTTATTTTAAATGTATTTGTTA  
CATTTTTAGTATTGATGTTGGGATAGGCATTTAAGCAAGTCTATAACTCACCTACATGCA  
[T, C]  
AATTTTGCCTTAATCAGTTTAAAGCTTTCTCTTAAATGAGAGATTTGAAATTCATAATTT  
CTGTGGTTCTTATCAGTTCTGAGTTTATTTTTTGGCCTTTTTATTTTTTAAAGGAAAA  
ATTGAGGCTTCAGAAATTGTCCAGTCTCTCCAGACACTGGGTCTGACTATTTCTGAACAA  
CAAGCAGAGTTGATTCTTCAAAGGTAAGCTCTTATGTTGGTCAACAATTGACTTTCACT  
TTAATATCCTGCATTAGAAGTCTGTGTTGTAAAGTGTGGCTTTAAACACCTCCCTAGTC

20941 GAGTTAGAATATCTTGATTTGGATTTAGTTCTGCTACTTCTTAATTACATTACTTGGTA  
AGGCCACTTGTGAAGTCACTCTTTTGGAGGAATATTATTTATCTATAAGGCTGTTACAA  
TTACTGAATTTTAAAAATGTGTATTTATTTTTTAAATGTATTTGTTACATTTTGTATT  
GATGTTGGGATAGGCATTTAAGCAAGTCTATAACTCACCTACATGCATAATTTTGCCTTA  
ATCAGTTTAAAGCTTTCTCTTAAATGAGAGATTTGAAATTCATAATTTCTGTGGTTCTTA

FIGURE 3, page 31 of 42

[T, C]  
CAGTTCTGAGTTTTATTTTTTGCCCTTTTTATTTTTTTAAAGGAAAAATTGAGGCTTCAG  
AAATTGTCCAGTCTCTCCAGACACTGGGTCTGACTATTTCTGAACAACAAGCAGAGTTGA  
TTCTTCAAAGGTAAGCTCTTCATGTTGGTCAACAATTGACTTTCACTTTAATATCCTGCA  
TTAGAACTCTGTGTTTGAAGTGTGGCTTTAAACACCTCCCTAGTCTTCATTATGTATA  
TCCAAGATCTTTTTGTCTTTTTTCTCCCATTCATTTGTATGTGTACATTTATCTAAAG

21116 GTATTGATGTTGGGATAGGCATTTAAGCAAGTCTATAACTCACCTACATGCATAATTTTG  
CCTTAATCAGTTTTAAAGCTTTCTCTTAAATGAGAGATTTGAAATTCATAATTTCTGTGGT  
TCTTATCAGTTCTGAGTTTTATTTTTTGCCCTTTTTATTTTTTTAAAGGAAAAATTGAGG  
CTTCAGAAATTGTCCAGTCTCTCCAGACACTGGGTCTGACTATTTCTGAACAACAAGCAG  
AGTTGATTCTTCAAAGGTAAGCTCTTCATGTTGGTCAACAATTGACTTTCACTTTAATAT  
[C, T]  
CTGCATTAGAACTCTGTGTTTGAAGTGTGGCTTTAAACACCTCCCTAGTCTTCATTAT  
GTATATCCAAGATCTTTTTGTCTTTTTTCTCCCATTCATTTGTATGTGTACATTTATC  
TAAAGTGTAAGAATGGGAAGTGTAAAGCTCAGACTGGACTCTTTCTTTCAAGGCCCTCAAAG  
GATAGTGGAATGGCAGGAAGTAAGGTTTTAACTCCATAGATGAGGAGCTGAAGAGTTTTG  
GTGTGCTTTTTCTCCATTTGATTCTAATGTGACAGTAAACTCATTGATTCAAACATAA

21701 CATTGATTCAAACTAAGAAGACTAGCAGATTCATCACATTATTTAACTAGATGTGACTG  
GAAAAAAGGGAAATTACTAAGCTCTCCAAGCTAACAAAGAAATACCTGTTTAACTTTCA  
GAAACAGAAATGCAAAATTTGAACCTTATTGTCTGGGGCAATCAGTTTGACTATTTAAGT  
CAGACTTTTATACTCTTAATGTTTGTGTTTCATGGGATAGAGCAGTAATCTCTGCAGCCCA  
GGTGCTCTCAAATACTCTGTTGCTATAAACACAGGGCAGGAAGTATTTTTTATGATAAC  
[G, A]  
TAAACAGAAAAGGACAATTATATTGTATTAATATTGTTGTGAATATTTTCAGTCCTCAC  
ATTGTCTAAAAATCTTTCTAAATGGCTTTGTTATTGAATTTATCTCATTTTATATCTGTG  
CCAACAGCATTTCATCCTTTCTCTTCATAATTTCTTTTACAAACAGCTGCTCAAGAGGA  
AGGCTCAAAGTCTCAAGGCTGAGCACGTAATGACTTTTGTTAGTACTAGATGAGAAGGGC  
TTTCCTGAGGAAATGAAAACCTAAACATGAAAAGAAGATAAACAGAATTTGGACAGTGA

21710 AAACCTAAGAAGACTAGCAGATTCATCACATTATTTAACTAGATGTGACTGGAAAAAAGG  
GAAATTACTAAGCTCTCCAAGCTAACAAAGAAATACCTGTTTAACTTTTCAAGAAACAGA  
AATGCAAAATTTGAACCTTATTGTCTGGGGCAATCAGTTTGACTATTTAAGTCAGACTTTT  
ATACTCTTAATGTTTGTGTTTCATGGGATAGAGCAGTAATCTCTGCAGCCCAGGTGCTCTC  
AAATACTCTGTTGCTATAAACACAGGGCAGGAAGTATTTTTTATGATAACGTAAACAG  
[A, -]  
AAAGGACAATTATATTGTATTAATATTGTTGTGAATATTTTCAGTCCTCACATTGTCTAA  
AAATCTTTCTAAATGGCTTTGTTATTGAATTTATCTCATTTTATATCTGTGCCAACAGCA  
TTTTTCATCCTTTCTCTTCATAATTTCTTTTACAAACAGCTGCTCAAGAGGAAGGCTCAAA  
GTCTCAAGGCTGAGCACGTAATGACTTTTGTTAGTACTAGATGAGAAGGGCTTTCTGAG  
GAAATGAAAACCTAAACATGAAAAGAAGATAAACAGAATTTGGACAGTGAGATATAGAG

21826 CAGAAATGCAAAATTTGAACCTTATTGTCTGGGGCAATCAGTTTGACTATTTAAGTCAGAC  
TTTTTATACTCTTAATGTTTGTGTTTCATGGGATAGAGCAGTAATCTCTGCAGCCCAGGTGC  
TCTCAAATACTCTGTTGCTATAAACACAGGGCAGGAAGTATTTTTTATGATAACGTAAA  
ACAGAAAAGGACAATTATATTGTATTAATATTGTTGTGAATATTTTCAGTCCTCACATTG  
TCTAAAAATCTTTCTAAATGGCTTTGTTATTGAATTTATCTCATTTTATATCTGTGCCAA  
[C, T]  
AGCATTTTCATCCTTTCTCTTCATAATTTCTTTTACAAACAGCTGCTCAAGAGGAAGGCT  
CAAAGTCTCAAGGCTGAGCACGTAATGACTTTTGTTAGTACTAGATGAGAAGGGCTTTCC  
TGAGGAAATGAAAACCTAAACATGAAAAGAAGATAAACAGAATTTGGACAGTGAGATAT  
AGAGCATATAATATTCTGCTTCTAAAGTAATATTCTTCTAGGAAAGTGAGGGCGTTTCCC  
TGCGTGTTAGGCCAGAAATCATATTCCTATATTTTCTTTGATAGCTTTAGGAATAATGCA

21840 TGAACCTTATTGTCTGGGGCAATCAGTTTGACTATTTAAGTCAGACTTTTATACTCTTAA  
TGTTTTGTTTCATGGGATAGAGCAGTAATCTCTGCAGCCCAGGTGCTCTCAAATACTCTG  
TTGCTATAAACACAGGGCAGGAAGTATTTTTTATGATAACGTAAAACAGAAAAGGACAA  
TTATATTGATTTAATATTGTTGTGAATATTTTCAGTCCTCACATTGTCTAAAAATCTTTT  
TAAATGGCTTTGTTATTGAATTTATCTCATTTTATATCTGTGCCAACAGCATTTCATCC  
[-, T]  
TTCTCTTCATAATTTCTTTTACAAACAGCTGCTCAAGAGGAAGGCTCAAAGTCTCAAGGC  
TGAGCACGTAATGACTTTTGTTAGTACTAGATGAGAAGGGCTTTCTGAGGAAATGAAAA

CCTAAACATGAAAAGAAGATAAACAGAATTTGGACAGTGAGATATAGAGCATATAATAT  
TCTGCTTCTAAAGTAATATTCTTCTAGGAAAGTGAGGGCGTTTCCCTGGCTGTTAGGCCA  
GAAATCATATTCTATATTTTCTTTGATAGCTTTAGGAATAATGCAAATCTAAGCCCA

21841 GAACCTTATTGTCTGGGGCAATCAGTTTGACTATTTAAGTCAGACTTTTATACTCTTAATG  
GTTTTGTTTCATGGGATAGAGCAGTAATCTCTGCAGCCCAGGTGCTCTCAAATACTCTGT  
TGCTATAAACACAGGGCAGGAAGTATTTTTATGATAACGTAAAACAGAAAAGGACAAT  
TATATTGTATTAATATTGTTGTGAATATTTTCAGTCTCACATTGTCTAAAAATCTTTCT  
AAATGGCTTTGTTATTGAATTTATCTCATTTTATATCTGTGCCAACAGCATTTTCATCCT  
[-, C, T]  
TCTCTTCATAATTTCTTTTACAAACAGCTGCTCAAGAGGAAGGCTCAAAGTCTCAAGGCT  
GAGCAGTAATGACTTTTGTAGTACTAGATGAGAAGGGCTTTCCTGAGGAAATGAAAAC  
CTAAACATGAAAAGAAGATAAACAGAATTTGGACAGTGAGATATAGAGCATATAATATT  
CTGCTTCTAAAGTAATATTCTTCTAGGAAAGTGAGGGCGTTTCCCTGGCTGTTAGGCCAG  
AAATCATATTCTATATTTTCTTTGATAGCTTTAGGAATAATGCAAATCTAAGCCCAAG

21843 ACCTTATTGTCTGGGGCAATCAGTTTGACTATTTAAGTCAGACTTTTATACTCTTAATGT  
TTTGTTCATGGGATAGAGCAGTAATCTCTGCAGCCCAGGTGCTCTCAAATACTCTGTG  
CTATAAACACAGGGCAGGAAGTATTTTTATGATAACGTAAAACAGAAAAGGACAATTA  
TATTGTATTAATATTGTTGTGAATATTTTCAGTCTCACATTGTCTAAAAATCTTTCTAA  
ATGGCTTTGTTATTGAATTTATCTCATTTTATATCTGTGCCAACAGCATTTTCATCCTTT  
[-, C]  
TCTTCATAATTTCTTTTACAAACAGCTGCTCAAGAGGAAGGCTCAAAGTCTCAAGGCTGA  
GCACGTAATGACTTTTGTAGTACTAGATGAGAAGGGCTTTCCTGAGGAAATGAAAACCT  
AAAACATGAAAAGAAGATAAACAGAATTTGGACAGTGAGATATAGAGCATATAATATTCT  
GCTTCTAAAGTAATATTCTTCTAGGAAAGTGAGGGCGTTTCCCTGGCTGTTAGGCCAGAA  
ATCATATTCTATATTTTCTTTGATAGCTTTAGGAATAATGCAAATCTAAGCCCAAGCT

22045 ATATTTTCAGTCTCACATTGTCTAAAAATCTTTCTAAATGGCTTTGTTATTGAATTTAT  
CTCATTTTATATCTGTGCCAACAGCATTTTCATCCTTTCTCTTCATAATTTCTTTTACAA  
ACAGCTGCTCAAGAGGAAGGCTCAAAGTCTCAAGGCTGAGCACGTAATGACTTTTGTAG  
TACTAGATGAGAAGGGCTTTCCTGAGGAAATGAAAACCTAAAACATGAAAAGAAGATAAA  
CAGAATTTGGACAGTGAGATATAGAGCATATAATATTCTGCTTCTAAAGTAATATTCTTC  
[C, A, T]  
AGGAAAGTGAGGGCGTTTCCCTGGCTGTTAGGCCAGAAATCATATTCTATATTTTCTTT  
GATAGCTTTAGGAATAATGCAAATCTAAGCCCAAGCTTCAGAATAGACTAAGAAGTATT  
AGCTTAGCTGCCATGACAAAATACCATAGGCTGGATGCATTAAACAATGGAAATTTAGTT  
TTTCACAGGTCTGGGAGCTGGGAAGTTTAAAGATGAGAGTGCCAGCATGGTTGGGTTGTAG  
TGAGGGCTCTCTTTCTGGCTTGACAGATAGACCCCTTCTCACTGTATTGTATATGGCAGA

22061 CATTTGTCTAAAAATCTTTCTAAATGGCTTTGTTATTGAATTTATCTCATTTTATATCTGT  
GCCAACAGCATTTTCATCCTTTCTCTTCATAATTTCTTTTACAAACAGCTGCTCAAGAGG  
AAGGCTCAAAGTCTCAAGGCTGAGCACGTAATGACTTTTGTAGTACTAGATGAGAAGGG  
CTTTCTGAGGAAATGAAAACCTAAAACATGAAAAGAAGATAAACAGAATTTGGACAGTG  
AGATATAGAGCATATAATATTCTGCTTCTAAAGTAATATTCTTCTAGGAAAGTGAGGGCG  
[G, T]  
TTCCCTGGCTGTTAGGCCAGAAATCATATTCTATATTTTCTTTGATAGCTTTAGGAATA  
ATGCAAATCTAAGCCCAAGCTTCAGAATAGACTAAGAAGTATTAGCTTAGCTGCCATGA  
CAAAATACCATAGGCTGGATGCATTAAACAATGGAAATTTAGTTTTTACAGGTCTGGGA  
GCTGGGAAGTTTAAAGATGAGAGTGCCAGCATGGTTGGGTTGTAGTGAGGGCTCTCTTTCT  
GGCTGCAGATAGACCCCTTCTCACTGTATTGTATATGGCAGAGAGAGAGAGAGAGA

22348 GAAAGTGAGGGCGTTTCCCTGGCTGTTAGGCCAGAAATCATATTCTATATTTTCTTTGA  
TAGCTTTAGGAATAATGCAAATCTAAGCCCAAGCTTCAGAATAGACTAAGAAGTATTAG  
CTTAGCTGCCATGACAAAATACCATAGGCTGGATGCATTAAACAATGGAAATTTAGTTTT  
TCACAGGTCTGGGAGCTGGGAAGTTTAAAGATGAGAGTGCCAGCATGGTTGGGTTGTAGTG  
AGGGCTCTCTTTCTGGCTTGACAGATAGACCCCTTCTCACTGTATTGTATATGGCAGAGA  
[-, A, G]  
AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGGGATCTTTCTCTTGCTTTCTATTATAAGG  
CCATAGTCTGTTGGATCAGGGTTCCATTCTTATGACTTTATTGACTTTACCCCCCTAA  
GATGCTATCTCCAGATATAATCACACGGTGGGTTAGGGCCTCAACATTTGGATTTGGGAG  
GGACACAGCTCAGTCCATAGCAAAGGATAATGCAGAGGGTTGGATATTTAAAAGTAGCTA  
CACAATTTTAAATATAAATATTTTATGGTAACTTTTTTTTTTTTTTTGAGATGGAGTCTAG

FIGURE 3, page 33 of 42





GGATATGCAATATAAAATATATAAAATTGTGACACTGAAAATTTAAAATGGGAGGAGTGGA  
GTAAAAGTACCTTCATATAACTTACTATTATATCTTATTGAATTGACCCTTTATCA  
TTATATAGGAACCTTGTCTCTCTTACAACCTCTGACTTAAAGTTTGTCTTATATGATA  
[T, C]  
AAGTAAAGTTACTCCTGCTCTCCTTTGGTTTCTGTTTCCATGGAATATCTTTTCCATTC  
CTTCACCATCAGTCTGTGTGTATTTTACAGATGAAATGAGTCTGTCTATGGGCAGCATAT  
AGTTGGATCTAGTTTTTTTAAATCCACTCAGACACTGTGTTTTTTGATTGGATAATTTAAT  
CCATTCATGTTCAAGGTAATTATTGATAAGTAAGGACTTTGTACTACCATTTTGCTTATT  
GTTTCATGGTTCTTTTATAGATCCTTTATTCTTTCTTCTCTCTGCTGTCTTTTTTTT

26060 GGTTTTTGGTTTGTGGTTACCAAGAGGTTACAAAAACATCTTAAGAGTTATAATAGTTT  
ATTTTAACTTGATAAATAATTTTATTGCAAAAACCCCCAAAAACAAAAAATCTACAC  
TTTTACTTAATCCCTGAAATTTTGAATTTTGTGATGTACAGTTTACCTCTTTTCATATT  
GTGTATCCCTTAAATTATTGTAGCTATTATTACTTTTAAATAGTTTCTCTTCTCTACTAC  
AGATGTAAGTGATTTGCATACCATCATTACAGTATTATTTTGAATTTACCTGTGTACTTT  
[C, T]  
TTTTATCAGCCAGTTTATACCTTTTCTGATGTTTTTGTGTTACTCATTAGCATCTTTTTCT  
TTCAGCTTGAGGAGCTCCTTTTACGTTTCTTATAAAATAGGTGCGGTGATGATTATCTCC  
CTCAGCTATTGTTGTCTGGGAAAGTATCTCTCTTCTCATTCTGAAGGACACTTTGCTGG  
GTACATTACCCTTGGTTGGTATTTTTCTCTTGAACGCTTTAAATATATCATCCCTTTCT  
CTCTGACCTGTTAGGTCTCTGCTGACCACTGTTTCCAACCATATTGGGACTGTCTTA

30245 ATTTTAAACCATCCATTGTTTCTGCTTCTCTAGATAACCTGACTAATATATAATTGGTAT  
GAAGTGATATCTCATGGCTTGTATTTATATTCTTTCATGGCTAGTGACTTTTTTGTAC  
TTTTGGGATATTGTTATTATTATTATTATTATTACTAGTGTTTTACTTCTTCAGTAAAA  
GTGTTAGAAACAATTTTTAAAGGCAGAATGTGACCAGAGTTTCTGTAGTTATATAACCA  
TCATGGACCTTCCCTCAAGTGCTAAGCCATTAGTGTTACTCATGTCACTCCAAATGTCAG  
[C, G]  
TTGTTTTCTTCCATTTCACTGTCTCTTTGTGTCCCAAACCTGAATTCATGGGAAAAACAT  
CTGAATGGTGCTTAATATGGTTTGGATATTGTCCCTCCAAATCTCATGTTGAAATATG  
ACCTCCAGTGTGGAAGTAGGACTACTTGGGTGACGAGAGTGATCCTTCATTAATGGC  
TTGGTAATAAGTGAACCTTATTAGTTTATGAAAGCTGGTTGTTGATAAGAGCCTGGCATC  
TCATTTCTCTGTCTCTCTCTCACCATTCTGACACACTTGCTCACCTTTTTTCTTCAGCCA

33664 TTCCAGAGTGTAGAAGTACACTGTCTATCCTTTCTAGGAGATCATTATAACACCAAAAG  
CAGACAGTATATGAAACAGGGAAATTAGAGGCCAAGATACCTATGACTTATATGTAAAAA  
TTTAAAGAAAATATTAGCAAATGAATCAGCCATTTTAAAAAATATACCACAATCAATGC  
ATTCATAAGAGCAGCTTAACAAAATTTGTTAGAAGGCATTAAAGAAGACTCAGTATAGAA  
AAGATGTACCTTCTCTCAAATTTGGTGATAGAGATTCAATGCCATTAAAAAACCCACCT  
[G, T]  
GTTTTTTTGGGAACTTGTCAAGCTGAGTCTCAAATTTATATCAAAGAGCAAAGGCCTAA  
GAATATCCAGGACATTCTGAAGAACTGTAAGGAGCCAGGGGCTGCCCTATCAGATACC  
AAGGGTTGTTATTAAGCCATAACCAAGTCAGTGCTGTTTCTACAGAAACAGACAAGTTAA  
CAAGTGAACATAATAGAGAGCCAGAAACAGACCCATCCATATTTGGATTTGTACGT  
GAAAGAAGTAGCTTTGCAAACTTTGGGAAAAGGAGAGTGTGTGCAATAGATGATGCTCG

33883 TAAAGAAGACTCAGTATAGAAAAGATGTACCTTCTCTCAAATTTGGTGATAGAGATTCAA  
TGCCATTAAAAAACCCACCTGGTTTTTTTTGAGGAACTTGTCAAGCTGAGTCTCAAATTT  
ATATCAAAGAGCAAAGGCCTAAGAATATCCAGGACATTCTGAAGAACTGTAAGGAGCCA  
GGGGCTGCCCTATCAGATACCAAGGGTTGTTATTAAGCCATAACCAAGTCAGTGCTGTT  
TCTACAGAAACAGACAAGTTAACAAGTGAACATAATAGAGAGCCAGAAACAGACCCAT  
[C, A]  
CATATTTTGGATTGTGACGTGAAAGAAGTAGCTTTGCAAACTTTGGGAAAAGGAGAGT  
GTGTGCAATAGATGATGCTCGTCTCATGCAGACAAAAGGAAATTGGGATACCTGCCTC  
TTACCGTACACAAACACCAACCTAAACGTGAAAGTTAAACTATAACAGCTTGAGTGCTG  
GGGAAGAAATATCTTTATCTCAGTGTAAGGAAGAATTTATTTTAAAAAGAAGACACAAAA  
GGCCATACATAGGAATGAAAAGATTGAATTCAGCTGCATTAAAAAGATTAAATTCAGCTG

34373 TATCTTTATCTCAGTGTAAGGAAGAATTTATTTTAAAAAGAAGACACAAAAGGCCATACA  
TAGGAATGAAAAGATTGAATTCAGCTGCATTAAAAAGATTAAATTCAGCTGCGTTAAAT  
CAAGAGCATCTGACTTGGACAGCATAGAGTGGAAGACAAAGAGAAGGTATTTGCCAGC  
TTATAACTTGAGGATTAGAATGAATGATATAAAGAAGTATGTAATAAGAAAAAGACAT  
ACAACCGGTTAGAAAAACGGGCAAGACATGAACAGCATATTTACGTGAAGGAAACAGC

[G, A]  
GTAGCAAATGAACATGGTAAGAGATGCTCAACACGTTTAGTAATTTGAAGGGAAATGCAA  
GTTATACCCACAGCAAGACTATCTTATCTAGGAAGTTTGTCAATACCCCTAAATGTTCTGT  
GGTTTAAAGCTACAGAGTTTGTAAATTCATTTATTTATTCATAAATACTCAGTGGCAGGC  
ACTGTTTATAGAAACCTTGGTTATAACTTTGAATGAAATTAATAAATACTCCTTGCCTTGTG  
GAGGATGCTTATGTGTGGGGAGTTGGGTGGTGGGTCAAACAACAATTACATTAAATAG

34558 ACTTGAAGGATTAGAATGAATGATATAAAGAACTATGTAAATAAGAAAAAGACATACAAC  
CGGTTAGAAAAACGGGCAAAGACATGAACAGCATATTTACGTGAAGGAAACAGCGGTAG  
CAAATGAACATGGTAAGAGATGCTCAACACGTTTAGTAATTTGAAGGGAAATGCAAGTTA  
TACCCACAGCAAGACTATCTTATCTAGGAAGTTTGTCAATACCCCTAAATGTTCTGTGGTT  
TTAAGCTACAGAGTTTGTAAATTCATTTATTTATTCATAAATACTCAGTGGCAGGCACTG  
[G, T]  
TTTAGAAACCTTGGTTATAACTTTGAATGAAATTAATAAATACTCCTTGCCTTGTGGAGGA  
TGCTTATGTGTGGGGAGTTGGGTGGTGGGGTCAAACAACAATTACATTAAATAGAAAAT  
AGTGACATAAATAAACCTATAAATATTGCAACCCAGAGTTATATTATAAATGTAAAGTAGT  
GACTAGGACTCTCATGCAGATATACCTCTGTGCTGGGACAAATGAAAGTTTAAAGTGTAAT  
TTCCCATATGCAAGTCAAAATAAAAAAGTGACACTAGAAAACACAATAATGAATATCTGAA

43929 GGCATTAAAGTATTCTGCCATAGGGAAGTGTAAGTTGTAGGCTTTTACTTTTTATAGG  
TACTATATTGTCCAAATAATCTCAGCACCTCATGGTTGCTAAGGATCTGTGTCTTGTGTT  
GGTCAGATTATGTTTATCTCTGGCATAAGGCACCTAACAAATATTCAATTAAAGGTTACAGA  
ATCTTTTTGCTTCATCTGCTTAGCATTTTCATACCAAGTTTGTGTTTCCACCAAACCTTCAAA  
TTTTGATTGTTTCATTAATATTCTGCATACTGATGTAAACCAAGTTCTATTATTGTGCAA  
[T, A]  
CTGCTCCTGAAACCTTAGGAACCTCTCTGAAGGAGTTTTATTTATTTTTGTTTTGTTT  
TTGTTTTGTTTTGTTTTTTTTGAGACGGAGTCTTGCTCTGTTGCCAGGCTAGAGTGCGAG  
TGGTGCGATCTCGGCTCTCTGCAAACTCGGCCTCCGGGGTTCACGCCATTCTCCTGCCTC  
AGCCACCGGAGTAGCTGGGACTACAGGCACCCACCACTGCGCCTGGCTAATTTTTTTTTGT  
ATTTTTAGTAGAGACGGGGTTTACCCTGTTAGCCAGGATGGTCTCGATCTCCTGACCTT

44309 TTGAGACGGAGTCTTGCTCTGTTGCCAGGCTAGAGTGCAAGTGGTGGATCTCGGCTCTC  
TGCAAACTCGGCCTCCGGGGTTTACGCCATTCTCCTGCCTCAGCCACCGGAGTAGCTGGG  
ACTACAGGCACCCACCACTGCGCCTGGCTAATTTTTTTTTGTATTTTTAGTAGAGACGGGG  
TTTCACCGTGTAGCCAGGATGGTCTCGATCTCCTGACCTTGTAAATCCGCCCGCCTCGCC  
TCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGTGCCCGGCCTTTTTTTTTTTTTTTT  
[T, -, C]  
TTTATGGGCTTGTCTTCTACACTTCAGATTTGACTAAATTAATATGCATTAAATGAAGT  
CAGGAGTTTACATTGCCACTAGTAACAATGCCCTAAGCTTACATAAAGCATTATAAAATTG  
TTGGTGATTAGTGCCCTTCTCAGCTATGAGTATAAGATAATATTACTAGTAGTTCAGTT  
GCCTAGATAAATGTACACTATGTGAAGTTTTATTTACATAATTCTTACGGTATTTTTTA  
AGGTAGTTGATAACAGTTGAGACTACAATTGTATCTCCATTTTATTGATAGTAAATGAA

44997 GAATTGTAAAAATATTATTATAGAATTTCTCTCAAACTATAGTAATGTAGAATAGGT  
TGAAGGGGTGATGATTTGAAACAATACCTCTCCATTAGCTAAATTTATATAGAATCTAT  
TGCGATGTTTTAAATGATAAGTCAGATTTATAAAAAATTTTTTATAAACAGTAGGAATGA  
GTTTAGGGGTATTCACATACAGTTTAAATTTTTATTTACATATTTAAACATATCATGGT  
ATAAATATGATGTGGATATAAATTTGAGATAAAGGAAGTATTGTTTAAGAATTGATGAAC  
[T, G]  
AATTTCTTAAAGATGTCATCACCAGTTGGTTTTCTAGCCTTATGAAAAATGGTTGCAAT  
AAAAAGATTGACTATGATAAAATGCTGCCCTTTCATTTTAACTAGACCAAGAGAAAAAC  
ATACTGTGAATCTATGATGAATGAAAGAAAGTTGTAAGTGTGGTTTGTATATTGTAA  
TTACTGTTTATTTTCATTTCTTGTGAACTGATACTGTACTTTGTTTCATTGTGAGTAGACA  
ACTTATAATCTATGTAATCAAATTTGGTTTAGTATAAATCTAGGGAATGAAGTTCATATT

46538 TGTTATACTTATGGTCAACACTTTTTATATTTGTCTGTAGATTTCTGTACAAAAGATTC  
TGACACTGTTTTAAGCCAGCATTCCTTCAGAATGTACCCAAATCTCAAAATTTATTTAGG  
GGCAAAGCTAATGCTTTAAAGAAAAAGGAGA  
[A, G]  
GGGATTGGTGTGTGTTTTCTTTAGGAACAGTAGTAAGTTGACTTTTAGAGAAGTTGAAT  
AAGCATTTATTTTTCTTTTGTCTTATTTATTTGTGAAGTTTATTTATTTAAATAAAAAT  
GGATTTCTCTGGAATTTAGTTTCTGCAATTTGAGGAGTTTCCAAAGTCAACCTTCAGGT  
TTGATACTTCTCTAGAAAGACTCACATAACTCACTGAAAGCTTATTACCCCTGGTTATGG

TTTATTACGGGGAAAAGATGCGGATGAAAATCAGTCAAGTAAAGAAGCACATAGGGCAGA

48153 TTATATCATTCTGCTTTTATTTTATAGGTTACGGTTCAAAATCAGACAAAATGAACATAT  
TTGGTGGCTTTTCGACAGATGGTAAAAGAAGGAGGTATCCGCTCGCTTTGGAGGGGAAATG  
GTACAAACGTCATCAAATTTGCTCCTGAGACAGCTGTTAAATTTCTGGGCATATGAACAGG  
TAATTGTTATCACCCGTGGAATTTATTAACAAAGAGGAGTTAGTAAACGGATTCAATAAA  
TGTTAATGTATAATGCTTTTGGGATTCTTGTTTAAATACATGATAATCTTTCACATATAC  
[T, C]  
CCATAAGGAGGATCACTTATAGGAGATTAGACTAAATAAAAATCAGAGATTTCTCATGACC  
AAGTTATGGGATTCTTAATTCATCATATTATTTATAAAGTTTTTTTTTTCTAAGTAGTTC  
TTAAAGGAAGGGTAGAATTTTAGTTTATTCTGAATCCTGAGCAGAAGCAGCACACT  
AACATAAGTTTTTATGAAAGTGTCACAATCTAACCTCTGGAAGGAAAACATAAGTTGAAG  
TCCTTTGTGTAATTTGACGTTGCTGTAAAATTGAGCTGAGTTTGGAGTGACACCTCCATG

48288 AAATTGCTCCTGAGACAGCTGTTAAATTTCTGGGCATATGAACAGGTAATTGTTATCACCC  
GTGGAATTTATTAAACAAAGAGGAGTTAGTAAACGGATTCAATAAATGTTAATGTATAATG  
CTTTTGGGATTCTTGTTTAAATACATGATAATCTTTCACATATACCCCATAGGAGGATC  
ACTTATAGGAGATTAGACTAAATAAAAATCAGAGATTTCTCATGACCAAGTTATGGGATTCT  
TTAATTCATCATATTATTTATAAAGTTTTTTTTTTCTAAGTAGTTCTTAAAGGAAGGGTA  
[G, T]  
AATTTTAGTTTATTCTGATCCTGAGCAGAAGCAGCACACTAACATAAGTTTTATG  
AAAGTGTCACAATCTAACCTCTGGAAGGAAAACATAAGTTGAAGTCTTTGTGTAATTT  
GACGTTGCTGTAAAATTGAGCTGAGTTTGGAGTGACACCTCCATGAAGGCAGGGGCGTGG  
CTTCTTCCCCATGTACTCCAGCACCTAGACAGAGCTTGGCATGTGATAAGTTTCAAGCGA  
GTGTTGAATGAGTCAATGAATGAACAAATGCATTTACCTCTGAATCACTTCTCTGTGCGG

48412 TGGGATTCTTGTTTTAATACATGATAATCTTTCACATATACCCCATAGGAGGATCACTT  
ATAGGAGATTAGACTAAATAAAAATCAGAGATTTCTCATGACCAAGTTATGGGATTCTTAA  
TTCATCATATTATTTATAAAGTTTTTTTTTTCTAAGTAGTTCTTAAAGGAAGGGTAGAAT  
TTTAGTTTATTCTGATCCTGAGCAGAAGCAGCACACTAACATAAGTTTTATGAAA  
GTGTCACAATCTAACCTCTGGAAGGAAAACATAAGTTGAAGTCTTTGTGTAATTTGAC  
[G, A]  
TTGCTGTAAAATTGAGCTGAGTTTGGAGTGACACCTCCATGAAGGCAGGGGCGTGGCTTC  
TTCCCCATGTACTCCAGCACCTAGACAGAGCTTGGCATGTGATAAGTTTCAAGCGAGTGT  
TGAAATGAGTCAATGAATGAACAAATGCATTTACCTCTGAATCACTTCTCTGTGCGGCTTT  
GTTAACTTGGATTATTTGAGCTATTGCTTCAGCCTAACTCAATGTAAGGGGAAATACAG  
AGGTAAGTTTTAGAGTTTGGGTTCTCTTATGGTCATTAGCAGAACTGTCTAGTTGAGCA

48446 CATATACCCCATAGGAGGATCACTTATAGGAGATTAGACTAAATAAAAATCAGAGATTTCT  
TCATGACCAAGTTATGGGATTCTTAATTCATCATATTATTTATAAAGTTTTTTTTTTCTA  
AGTAGTTCTTAAAGGAAGGGTAGAATTTTAGTTTATTCTGATCCTGAGCAGAAGC  
AGCACACTAACATAAGTTTTATGAAAGTGTCACAATCTAACCTCTGGAAGGAAAACATA  
AGTTGAAGTCTTTGTGTAATTTGACGTTGCTGTAAAATTGAGCTGAGTTTGGAGTGACA  
[C, G]  
CTCCATGAAGGCAGGGGCGTGGCTTCTTCCCCATGTACTCCAGCACCTAGACAGAGCTTG  
GCATGTGATAAGTTTCAAGCGAGTGTGTAATGAGTCAATGAATGAACAAATGCATTTACC  
TCTGAATCACTTCTCTGTGCGGCTTTTGTAACTTGGATTATTTGAGCTATTGCTTCAGCC  
TAACTCAATGTAAGGGGAAATACAGAGGTAAGTTTTAGAGTTTGGGTTCTCTTTATGGT  
CATTAGCAGAACTGTCTAGTTGAGCAGCCACAGATTATGTTTTCCATTATTTATTCATC

48456 ATAAGGAGGATCACTTATAGGAGATTAGACTAAATAAAAATCAGAGATTTCTCATGACCAA  
GTTATGGGATTCTTAATTCATCATATTATTTATAAAGTTTTTTTTTTCTAAGTAGTTCTT  
AAAGGAAGGGTAGAATTTTAGTTTATTCTGATCCTGAGCAGAAGCAGCACACTAA  
CATAAGTTTTATGAAAGTGTCACAATCTAACCTCTGGAAGGAAAACATAAGTTGAAGTC  
CTTTGTGTAATTTGACGTTGCTGTAAAATTGAGCTGAGTTTGGAGTGACACCTCCATGAA  
[G, C]  
GCAGGGGCGTGGCTTCTTCCCCATGTACTCCAGCACCTAGACAGAGCTTGGCATGTGATA  
AGTTTCAAGCGAGTGTGTAATGAGTCAATGAATGAACAAATGCATTTACCTCTGAATCAC  
TTCTCTGTGCGGCTTTTGTAACTTGGATTATTTGAGCTATTGCTTCAGCCTAACTCAATG  
TAAAGGGGAAATACAGAGGTAAGTTTTAGAGTTTGGGTTCTCTTTATGGTCATTAGCAGA  
ACTGTCTAGTTGAGCAGCCACAGATTATGTTTTCCATTATTTATTCATCATTGTTTATC

48789 GCACCTAGACAGAGCTTGGCATGTGATAAGTTTCAAGCGAGTGTGAAATGAGTCAATGAA

TGAACAAATGCATTTACCTCTGAATCACTTCTCTGTCTGCGCTTTTGTAACTTGGATTATT  
TGAGCTATTGCTTCAGCCTAACTCAATGTAAAGGGGAAATACAGAGGTAAGTTTTAGAGT  
TTGGGTTCTCTTTATGGTCATTAGCAGAACTGTCTAGTTGAGCAGCCACAGATTATGTTT  
TCCATTATTTATCCATCATTGTTTATCAAGGACTGTAAGGGCCTTGAAATTCAACTCCC  
[C, -]  
CCCCCATAGTTTTTGTATTATTCATGTAGATTTTAGATTATTCTGGAGAGTGTTTTGTT  
CTTGAGCAACAGAATACTCTTGAGAAGATTACGAAGTCCAGTGGTATCCTTTTCTTTGCC  
TAGGAAATAGAGAAGCAAAAAAAAAAAAAAAAAAATAAAGAAAATCTAGTCTCCAGG  
ATTTTAATTAGAACCTATCCTTGGGAAGGCTATTTTCTTATATGAAGGTTTGAAGATTC  
AAATCATGATTATTAAGGGCTAATGTTTGTAGATACCCTTAGGTTATTCTGACCACATACT

48859 CATTACCTCTGAATCACTTCTCTGTCTGCGCTTTTGTAACTTGGATTATTGAGCTATTG  
CTTCAGCCTAACTCAATGTAAAGGGGAAATACAGAGGTAAGTTTTAGAGTTTGGGTTCTC  
TTTATGGTCATTAGCAGAACTGTCTAGTTGAGCAGCCACAGATTATGTTTTCCATTATTT  
ATTCATCATTGTTTATCAAGGACTGTAAGGGCCTTGAAATTCAACTCCCCCCCCCATAG  
TTTTTGTATTATTCATGTAGATTTTAGATTATTCTGGAGAGTGTTTTGTTCTTGAGCAA  
[G, C]  
AGAATACTCTTGAGAAGATTACGAAGTCCAGTGGTATCCTTTTCTTTGCCTAGGAAATAG  
AGAAGCAAAAAAAAAAAAAAAAAAATAAAGAAAATCTAGTCTCCAGGATTTTAAATTA  
GAACCTATCCTTGGGAAGGCTATTTTCTTATATGAAGGTTTGAAGATTCAAATCATGAT  
TATTAAGGGCTAATGTTTGTAGATACCCTTAGGTTATTCTGACCACATACTTGGATTTTAT  
GATAGGAAAGCCACAGCCTAAATAAATAAATACTCAATGCAGTTATTTTCAAGTATGCAAG

49126 GATTATTCTGGAGAGTGTTTTGTTCTTGAGCAACAGAATACTCTTGAGAAGATTACGAAG  
TCCAGTGGTATCCTTTTCTTTGCCTAGGAAATAGAGAAGCAAAAAAAAAAAAAAAAAAAAA  
ATTAAGAAAATCTAGTCTCCAGGATTTTAAATTAGAACCTATCCTTGGGAAGGCTATTTT  
CCTTATATGAAGGTTTGAAGATTCAAATCATGATTATTAAGGGCTAATGTTTGTAGATACC  
CTTAGGTTATTCTGACCACATACTTGGATTTTATGATAGGAAAGCCACAGCCTAAATAA  
[A, G]  
TAAATACTCAATGCAGTTATTTTCAAGTATGCAAGAAGTTTGGTATTTTGTAAAAAGTCCAT  
GGGTATTGCAAGCAAATATGCACATTTTGCTTTATGCCATTTGTGAGATTCTTACCTTGG  
ATACCACCAACAGGCATCCTCTGTCTCCACCAAGCTCCTTCTGAGACCTCTTTATAGTATTGTGA  
TAGTATTGTGATTCTGACACTAACTTTTCTTAGACATGAAGAGAAAGCTGTCTACACAG  
TGTGGTGTAGTTTTCTTATGGGCTCTGGACCTATGGTGCTGTTTTCTCTCCTCTGCTGA

49378 TGACCACATACTTGGATTTTATGATAGGAAAGCCACAGCCTAAATAAATAAATACTCAA  
TGCAGTTATTTTCAAGTATGCAAGAAGTTTGGTATTTTGTAAAAAGTCCATGGGTATTGCAA  
GCAAATATGCACATTTTGCTTTATGCCATTTGTGAGATTCTTACCTTGGATACCACCAAC  
AGGCATCCTCTGCTTCTGTCCACCAAGCTCCTTCTGAGACCTCTTTATAGTATTGTGA  
TTTCTGACACTAACTTTTCTTAGACATGAAGAGAAAGCTGTCTACACAGTGTGGTGTAGT  
[T, G]  
TTCTTATGGGCTCTGGACCTATGGTGCTGTTTTCTCTCCTCTGCTGAAGGTCCATTCTAT  
CCCTCGGGGCTCTCTAAAAGCCACCTTCTGTGACAAGCATATACTAAGCATCTCAATCA  
AAGCCAGTTCTCTCCCTGTCCAGCCTCCCTCGAGTGCTGAATTGCAGAATATCCCATTTT  
TCATTGGATGATGGAAGCCATTGTTTTCCAGTGGATTGTAAATTACTTCGGGGTAA  
TAGGCTGTATATATTCTCAAATTTCCAGAGTATGTAAGTCAAGTCACTTTTAGATTGAGA

49482 TCCATGGGTATTGCAAGCAAATATGCACATTTTGCTTTATGCCATTTGTGAGATTCTTAC  
CTTGATACCACCAACAGGCATCCTCTGCTTCTGTCCACCAAGCTCCTTCTGAGACCT  
CTTTATAGTATTGTGATTTCTGACACTAACTTTCTTAGACATGAAGAGAAAGCTGTCTA  
CACAGTGTGGTGTAGTTTTCTTATGGGCTCTGGACCTATGGTGCTGTTTTCTCTCCTCT  
GCTGAAGGTCCATTCTCCCTCGGGGCTCTCTAAAAGCCACCTTCTGTGACAAGCATAT  
[A, C]  
CTAAGCATCTCAATCAAAGCCAGTTCTCTCCCTGTCCAGCCTCCCTCGAGTGCTGAATTG  
CAGAATATCCCATTTTTCATTGGATGATGGAAGCCATTGTTTTCCAGTGGATTGTAA  
ATTACTTCGGGGTAAATAGGCTGTATATATTCTCAAATTTCCAGAGTATGTAAGTAGGT  
CACTTTTAGATTGAGATAGATTTTGTTCCTTGAATAGCTAGTACTTTAGGAACTAAGAA  
AAAGATCTTTTCAACCTGGTATGTAGCTCTGTCAAACACATCATCAGTATGGGGTAAACC

49741 CTCGGGGCTCTCTAAAAGCCACCTTCTGTGACAAGCATATACTAAGCATCTCAATCAAA  
GCCAGTTCTCTCCCTGTCCAGCCTCCCTCGAGTGCTGAATTGCAGAATATCCCATTTTTC  
ATTGGATGATGGAAGCCATTGTTTTCCAGTGGATTGTAAATTACTTCGGGGTAAATA  
GGCTGTATATATTCTCAAATTTCCAGAGTATGTAAGTCAAGTCACTTTTAGATTGAGATA

GATTTTGTTCCTTGAATAGCTAGTACTTTAGGAACTAAGAAAAAGATCTTTTCAACCTG  
[G, A]  
TATGTAGCTCTGTCAAACACATCATCAGTATGGGGTAAACCTGTGTTCTCTGTGGGTGT  
CATTACCATAGTAGTGTCTATTGTATCATTGACAGTGTAATAGTGTGGGGTAGTGTCTTG  
TGGTTTCAGCTGCCACTCTGTACTGACTGCTTTCCACTCCA

49840 ATCTTTTCAACCTGGTATGTAGCTCTGTCAAACACATCATCAGTATGGGGTAAACCTGTG  
TTCTCTGTGGGTGTCTATTACCATAGTAGTGTCTATTGTATCATTGACAGTGTA  
[A, G]  
TAGTGTGGGGTAGTGTCTTGTGGTTTCAGCTGCCACTCTGTACTGACTGCTTTCCACTC  
CAACATCTTCTCTTTATCTCAACACTGTAGGTCTACCTGTGTACTGTGTGTTTCAGCAT  
CTCTGCTTGCATGACCCAGGAGTGCCTCCCACTCAATATGGCCACCATGCATGGTCATCT  
TTCTGCTACTCCCTGTCTCCTGACCCTGCTCCAGCAACACAGACAGACACCCTTCTCTT  
TCTATATGTCATATGGTGGGAATGCCCTTTAGTACTTACTCAGGAGTTAGTTCCTCTGG

50102 CATTACCATAGTAGTGTCTATTGTATCATTGACAGTGTAATAGTGTGGGGTAGTGTCTTG  
TGGTTTCAGCTGCCACTCTGTACTGACTGCTTTCCACTCCAACATCTTCTCTTTATCTC  
AACACTGTAGGTCTACCTGTGTACTGTGTGTTTCAGCATCTCTGCTTGCATGACCCAGGA  
GTGCTTCCCACTCAATATGGCCACCATGCATGGTCATCTTTCTGCTACTCCCTGTCTCCT  
GACCCTGCTCCAGCAACACAGACAGACACCCTTCTCTTTCTATATGTCATATGGTGGGG  
[G, A]  
ATGCCCTTTAGTACTTACTCAGGAGTTAGTTCCTCTGGGAAGCCTTCTGTTCTAGTTTCC  
TTTTGTTACAGCACTTTTACATTGAATTCTGACGTTCTCTGTACTTATCTGCTTTGTGAG  
ACTGTGAGCTTCTTAGGCAGTAGCTACTTGTATTCTTAGCACCTTGCCAGTGCCAGGA  
AACCTTTATTAAGTAAATGAAAAGACAGAACTGACAGACTGGAATTAGAGCTCAAGCTTG  
CCTCAATCTCAAGCCATTAAGATGAAGGGGAGCCGGGCGTGGTGGCTCACGCCTCTAATC

50109 ATAGTAGTGTCTATTGTATCATTGACAGTGTAATAGTGTGGGGTAGTGTCTTGTGGTTTC  
AGCTGCCACTCTGTACTGACTGCTTTCCACTCCAACATCTTCTCTTTATCTCAACACTG  
TAGGTCTACCTGTGTACTGTGTGTTTCAGCATCTCTGCTTGCATGACCCAGGAGTGCCCTC  
CCACTCAATATGGCCACCATGCATGGTCATCTTTCTGCTACTCCCTGTCTCCTGACCCTG  
CTCCAGCAACACAGACAGACACCCTTCTCTTTCTATATGTCATATGGTGGGGAATGCCC  
[C, G, T]  
TTAGTACTTACTCAGGAGTTAGTTCCTCTGGGAAGCCTTCTGTTCTAGTTTCTCTTTTGT  
ACAGCACTTTTACATTGAATTCTGACGTTCTCTGTACTTATCTGCTTTGTGAGACTGTGA  
GCTTCTTAGGCAGTAGCTACTTGTATTCTTAGCACCTTGCCAGTGCCAGGAACCCCTT  
ATTAAGTAAATGAAAAGACAGAACTGACAGACTGGAATTAGAGCTCAAGCTTGCCCTCAAT  
CTCAAGCCATTAAGATGAAGGGGAGCCGGGCGTGGTGGCTCACGCCTCTAATCCCAGCAC

50747 CCAGCTGGGCAACGTGGCAAAACCCCATTTCTACAAAAATATAAAAAATTAGTTGGACG  
TGGGGGTGTGTGCCGTACTCAGGATGCTGAGGTGGGAGGATCACTTGAGCTCGAGAGGC  
AGAGGTTGCAGTGAGCTGGGATCACACCATTGCAATCTAGCCTGGGTGATAGAATGAGAC  
CTTGCTCTAAAAAATAAATAAATAAAGGGGAAGATAAGGATTGGAACAGAA  
GGAGCAGCATGTGGACAGAAATGTAGGCACAAGAAGGCATCACTCACTGAAGAGACTGAA  
[G, A]  
GTGGTTCACTGTGCCTCAAGACTGGTGGAGTGTGTTTCCGGAAGATAATGATGAAAGAG  
CTGGACAGATAAACAGGGGCCAAATGTAATAGGAGTCTGGATTTTATTCTGAATATGGTA  
GGGGCTATTGTAGCATCTTATATAGGGAAGTGAATGAGTACATTACATTTAAGGAATA  
TCAACCTGAAAAAGAGTGGAGACATTGTTGGGGGAGAGTGAAGTAGACTAGAGGCAGGG  
AGAAATTTTAAATAATTGAGGTAAGAAATGATGAACACCAGTATAAGGTGATGTCTTTAA

51272 TAGACTAGAGGCAGGAGAAATATTTAAATAATTGAGGTAAGAAATGATGAACACCAGTAT  
AAGGTGATGTCTTTAAGGAATGGAGAAGGAATGAACTGAGAAATATTTTGAAGTAGAA  
TCAACAGAACTCACTGACTGACTGGATATGGAGGTGAGAAAGAGAAGAGTCAAGAATGAT  
ATTCTAATTTCTAAGTGTAGTGCATTCAAAGAGAATACAATATCAGGTTCCATTTT  
GTGCATGCTGAGTTTGTAGATGTGTGGGACATGTACAGGGAGCTGTCCAGTAAGCAATTGG  
[G, A]  
TATATCAGCTAGCCATTAAGAGAGAGATCTTTGATAGAGAGGTTGTGCTGAGTTGAGCC  
ATTGGAATGGGCAGGATCACTCAAGAAGAGCTTATAAATGAGAAGAATTCTAGGAATAAG  
TCCAAAGGGAGAAGTAAAAGAAGAACTTGCAAAGGACACTGAGAAGAAATAGCTCGAGG  
GATGGGAGAAAATCCAGAGAGAGGGATGGCATAGGAGTCAAGTGAAGGAAACGGTTTCAT  
GGGGGTCAGTACTACTGGGTAGTGAATATAATAAGAATATCTTTTAGGATTTCTCAACCC

52842 TCAGGGTGGTTTTGAGGGCTCAGTTAAGTCTCCTTTAGGAAGGTTGAGTTCTGTAGCCTT  
GGCAAGTTACTTTAAAGTCTCTGTGACTATTACCTCATCTCTAAGATGGGGACTAAGCTTG  
GTGACATAGTTTTACATACCAGGCACAGTGCCTGACTTTTTGGCTCTGTCTGAAAGTCTT  
CCCTTTGTATATGGTATGTTTCGGGGAATAGGAGCCTCAAGCACTTATCCTTTAAATATT  
TATCCTCCATCAGTCACTAAACGTTTACTCTGTACTTTTGATAGGTGCTGTGGGGCTCA  
[G, A]  
GGTATAAAAGGTACCTTCAAAGTTACTGTTAAAGTGCAGGAAGGTTTTTAAGCAAATTAT  
GTTTAATGATTTTGACAATCTGACATGCAGGAAAATTAATAGGGCCTATGCAGAAGAGGA  
GTTTTATGTAACACTCTGTAGTTGAGGAAACAGAGCCCTTGGAAAGCAGTGATCTCTCTGG  
GGAGGAATGTCTGGTATTTGGGAATCTCATGAAATGATAATATACTTAATTTTTATCATG  
AGCAGCAAAACACAGATTTGCTAGGAGAAAGTCATCGTATGTTGTTGCATTGGGCACTTT

61837 GAGGAACCTCCATGTCATTTTCCATAGTAAGTAGACCTTTTTGTTTTTAAACATTCTAT  
CAATGTACACCAAGATTCCAATTTCTCCATGTCCTCCCAACACCATTAAAGTGGGGTGGT  
GGTCTACTACTATTGCTGTGTGTGCTGTTTATTCTCCTCCCTTCAGTTCTGTAAGTGGTTGCT  
TCATATATTTAGGAGCTTAATATTAGGTCCATATGAAGTTATAATTTCTTCTGGTAAAG  
TGACCCATTTATCATTATGTAATGTCCATCTTTGTCTCTTGTGACAGTTTGTGTCTTAA  
[A, G]  
TCTATTTTGTCTGATGTAATTATGGCCACCCCTTTTCTCTTTGGGTTCCCGTTTTTATGG  
AATATCTTTTTCCATCCTTTCACTTTGAGCTTATGTGTGCTTAGATCTAAAGTGAGTC  
TCATAGATAAGGTATAGTTGATTCTGTATGTGTTATTTCACTCAGCAATTTATATCTTTTA  
GTTAGGGGATTTAATCCATTTACATTTAAAGCAGTTACTGATAGGGAAGGACTTACTGTT  
GTCATTTGGCTAGCTACCTTTTTATCTTTGTCTGTGGCTTTTCTGTTTTTCCCTTCCTC

62018 CATATATTTAGGAGCTTAATATTAGGTCCATATGAAGTTATAATTTCTTCTGGTAAAGT  
GACCCATTTATCATTATGTAATGTCCATCTTTGTCTCTTGTGACAGTTTGTGTCTTAAAA  
TCTATTTTGTCTGATGTAATTATGGCCACCCCTTTTCTCTTTGGGTTCCCGTTTTTATGG  
AATATCTTTTTCCATCCTTTCACTTTGAGCTTATGTGTGCTTAGATCTAAAGTGAGTC  
TCATAGATAAGGTATAGTTGATTCTGTATGTGTTATTTCACTCAGCAATTTATATCTTTTA  
[A, G]  
TTAGGGGATTTAATCCATTTACATTTAAAGCAGTTACTGATAGGGAAGGACTTACTGTTG  
TCATTTGGCTAGCTACCTTTTTATCTTTGTCTGTGGCTTTTCTGTTTTTCCCTTCCTCT  
CTTCTGGCTTCTTCTGTGTTTTGTTGATTTTTTTTTTTTTTTTGTAGTGATATGTTCTGAT  
TCCCTTCTCATTTCCCTTTGTGTGCATTCTATAGATGCTATTTTGTGGTTACCATTGCA  
ACTACATAAAGCATACTAAAGTTATAGCAACTTATTTAAGCTGTTTACAACCTAACTTC

65562 GACTGAAATTCAGACACATGCAGTCTGATTCTAACCCTCCTGTCTGCCAGCTCTGATCCA  
GAACTTTGCATGACTGATACGGCTGATAGATTGTCTATGGCTGATAGACTGTCATTTCTG  
ACCTAAAAGTCTGATCATTTTACATCTGTTGAGACATCTTTGAGCCCTTTGCGTGTGAGT  
TCCAAAGTTGTTAGTGGGAATTTCAAAGCCTTTAATAATCTAGCCCCACTTTGTTCACTC  
TCTGTGTAATAACCACATACAACAATTGGCTGCATCTCCATAGCACATGGTACTCCTCCC  
[A, G]  
TTGTCTTGGTTGTGCCAGCAACACTGGTTTTCGCTTTCTCTTCTGCTTGTGAGGTCAT  
TTCCAAGGCCAGGTCTTTGTGCTTTTTCCCAAGCTTCCCAGAGCTTCTTCCATACTCCC  
CTTACTTCTGAGATTTAAGTGTCTCTTCTCAGCGCTTGTCTAGTAAGAAGGAGGAGC  
AGCAGCACTGTGGGTGGTGGAAAGTGTACCAGCTTTGGAGTCAGACCATTGGATCTCAG  
CCCTACCATTTTCTACTTAGATTTTTTTAGGACAAATTTCTCCATCTTTCTAAGCCTCCA

65780 TCTAGCCCCACTTTGTTCACTCTCTGTGTAATAACCACATACAACAATTGGCTGCATCTC  
CATAGCACATGGTACTCCTCCCGTTGTCTTGGTTGTGCCAGCAACACTGGTTTTCGCTTT  
CTCTTCTGCTTGTGAGGTCAATTTCCAAGGCCAGGTCTTTGTGCTTTTTCCCAAGCTT  
CCCAGAGCTTCTCCATACTCCCTTACTTCTGAGATTTAACTGTTCTCTCTTCAAGCGC  
TTGTCTAGTAAGAAGGAGGAGCAGCAGCACTGTGGGTGGTGGAAAGTGTACCAGCTTT  
[G, A]  
GAGTCAGACCATTGGATCTCAGCCCTACCATTTTCTACTTAGATTTTTTTAGGACAAATT  
TCTCCATCTTTCTAAGCCTCCAATTGCTCACTTACAAAATTGATATAACATTTACCTTGC  
AAGATTGGTATGGAAGGTAATTAACCCAGTATTTAGAACATAGTAATTAATAAATAACTA  
TTATTACCATCATTACTATAGTTAGGACACTCACTGTTAGGTGCTATACAAAGAGGATCA  
TAAAAGGGATGTTGTCTTGGGCTTCTTGGAAATAAATGTTGTCTTTTACTGTATTTTGA

66092 TTGGATCTCAGCCCTACCATTTTCTACTTAGATTTTTTTAGGACAAATTTCTCCATCTTT  
CTAAGCCTCCAATTGCTCACTTACAAAATTGATATAACATTTACCTTGCAAGATTGGTAT  
GGAAGGTAATTAACCCAGTATTTAGAACATAGTAATTAATAAATAACTATTATTACCATC

FIGURE 3, page 40 of 42

ATTACTATAGTTAGGACACTCACTGTTAGGTGCTATACAAAGAGGATCATAAAAGGGATG  
TTGTCCTTGGGCTTCTTGAATAAATGTTGTCCTTTTACTGTATTTTAGAATATCATTCTG  
[G, A]  
GTCATAATTGTTTGTGTCATAATAATGAAACATACTTGAATATTAAATTACCCTCTTTT  
TTTATTTTTTAGCCATGTTAGAAGGTTCCCCACAGCTGAATATGGTTGGCCTCTTTTCGAC  
GAATTATTTCCAAAGAAGGAATACCAGGACTTTACAGAGGCATCACCCCAAACCTTCATGA  
AGGTGCTCCCTGCTGTAGGCATCAGTTATGTGGTTTATGAAAATATGAAGCAAACCTTAG  
GAGTAACCCAGAAATGATGTTGCATTTTTTGCTTTAGCCTGATAATTGAAACTTTCAACA

66617 ATGAAGCAAACCTTTAGGAGTAACCCAGAAATGATGTTGCATTTTTTGCTTTAGCCTGATA  
ATTGAAACCTTTCAACAATCTCTGGAGTGACTTTTTCTCCTCGAATTGAAACAAGTCTATG  
GCAAAGAAGCTGCATTTTTTTCACAAAAGGGAAGATGGTAACAATGGTCACTTCAAAC  
TTTGGGCTAAATTATATGTACACAGAAATGTTCAAATCATAGTTTAAATGTGTTTGAA  
AAGGCCACACAATTATACTTTATCTTTCTTAATAATCCTGCAATCTCTGCCCTGAATC  
[C, T]  
GAAATCTGAAATGTACTGGCTTGAACAAAATTTGTTTTGTGTGTTAGAGTTATAAATCA  
TTAATCTTTATTTTCGGGTGGTTTACGTTTATGCCAGTTCCCTTTATATTTAAATTTCTTGT  
TTTATATATTTGAATGTCTTTATAGATTTCTTTAAATTTCTTTATAGAACCATTAAATAG  
AAATCATTTACATTTAAATATACCTTACAGCAAAAGCATCCAAATAAGTATAGGGTTTA  
TGTCCTTATTTTTCTTTAGCTGAATACGAATGAGCACAGTGGTGAATTTCTGAAGGGA

66892 ATCTGCAATCTCTGCCCTGAATCCGAAATCTGAAATGTACTGGCTTGAACAAAATTT  
GTTTTGTGTGTTAGAGTTATAAATCATTAATCTTTATTTTCGGGTGGTTTACGTTTATGCC  
AGTTCCTTTATATTTAAATTTCTTGTTTTATATATTTGAATGTCTTTATAGATTTCTTT  
AAATTTCTTTATAGAACCATTAAATAGAAATCATTACATTTAAATATACCTTACAGCAA  
AAGCATCCAAATAAGTATAGGGTTTATGTCTTATTTTTCTTTTACGCTGAATACGAATGA  
[G, A]  
CACAGTGGTGGAATTTCTGAAGGGAAGTGATGAAATTATATTTATTTTCACTGGGCACCTT  
TCCATTTTACCCTGTACCATTATTTGGTTCCTGGAGTTATACACTAATTTTTCAGTATAT  
TACTGTTAAATTACCAACACAAGGCAATTTATTTGAAAGATTCCGTTTATCCTGCCATTG  
CTTTGAAAAGCAGCAGGAAACGAAATCCTTTGACTTGTATCAGCTTCTGCAGAGCATCTT  
TGTTTTCTTTTGTCTTTGTTTCTTACCTTTTGAATCAGATTCCGTTTTAGTCAGGAAGA

67263 CACTGTACCATTATTTGGTTCCTGGAGTTATACACTAATTTTCACTATATTACTGTTAA  
TTACCAACACAAGGCAATTTATTTGAAAGATTCCGTTTATCCTGCCATTGCTTTGAAAAG  
CAGCAGGAAACGAAATCCTTTGACTTGTATCAGCTTCTGCAGAGCATCTTTGTTTTCTT  
TGTCCTTTGTTTCTTACCTTTTGAATCAGATTCCGTTTTAGTCAGGAAGACTTCTTGGGA  
CCATTCTTAGTAACCTGAAATTTCTTTTTTAATGTCATGAAGTGGATTGATCATGAGCAA  
[G, A]  
TGATGTGCTTATTTCTCCCTCACTGTTGAATATCTTTGAACTTGCTGTTTTCAATATGGG  
CAGCACAAGGTGAGAGATACATATTAATAGTAGTATGTATTACTCTTATACATTAGATA  
CCTATATTTAAATGAAAGGCCCAATTTGTAAACATATACATTTCATTTCTCTCTTCCCC  
AAGTTTTAGGAACATGTTAGGATATAGGAGACTTAATTTATAATAATGAGAGCATTTTTT  
TATTTTACTAAAGCCATTTTATAGTCAACTATCTTTCTTATTTGTGTGATTAGAACTT

67651 ATAGTAGTATGTATTACTCTTATACATTAGATACCTATATTTAAATGAAAGGCCCAATTT  
GTAAACATATACATTTCATATTCTCTCTTGGCCCAAGTTTTAGGAACATGTTAGGATATAG  
GAGACTTAATTTATAATAATGAGAGCATTTTTTTTATTTTACTAAAGCCATTTTATAGTC  
AACTATCTTTTCTTATTTGTGTGATTAGAACTTAGAAAAATATTTACTAGTTGAAGTTAT  
TATCAGTTTTTTAATTTAGTTCTTAAACTCATTTCACTTCTAATAATTTCTGTTATAAAT  
[G, T]  
CCAGCATTTTAAATGAAAATCTAATGATGTAATAGGCATTTTCTTTATTTGAACTTACCTC  
TTTTATTTTCTGAACCAAAGAGAAAGATGGACTGGTGTGTTGTGAAACATTTTTAAATG  
TAGTTTCATTTATATTAGTTATGTTTGATAAATGTCTCAGTATTTTATAATATGATAAG  
CCTGGGATTCTACTTTTAGGGTTATTTGTACTTTTGTAGTAATATATAAAGTGACAAATAT  
AAGGTACATGATCAGCTCTTTCTATTTTTTACTCGTAAAAATATGGAAATGAATAATTTT

67935 ATTTCTGTTATAAATTGCCAGCATTTTAAATGAAAATCTAATGATGTAATAGGCATTTTCT  
TTATTTGAACCTACCTCTTTTATTTTCTGAACCAAAGAGAAAGATGGACTGGTGTGTTGTG  
AAACATTTTTAAATGTTAGTTTCATTTATATTAGTTATGTTTGATAAATGTCTCAGTAT  
TTTTATAATATGATAAGCCTGGGATTCTACTTTTAGGGTTATTTGTACTTTTGAGTAATA  
TATAAAGTGACAATATTAAGGTACATGATCAGCTCTTTCTATTTTTTACTCGTAAAAATTA  
[C, T]

FIGURE 3, page 41 of 42

GGAAATGAATAATTTTGCTAACAACTTTGAAATTTCAAACCTTCTGGAAAATATGAAAATA  
TTCATTGTTTCATTATGAATTTAAATTGTAAGGTATGAATGTGATTGTCTGTACATCTTG  
TATCTTTTCCAAAAATGATTCTGTATCTTTTGGAAAAAGCCGAGAGTTGAAGATAGTA  
TATTTCTGGTAGTACTGAATATTTACTTACAGTTTCTATCAAAAATATATATTTGTTTCT  
AAAATTACTTGTTCAGTTTTATTTTTTTTAGAGAAAATTCTTAAGTCTCAGTTTCC

69000

TTCAGAAATAACTTATCAGTTATTTCTGTAAGCTTCTTGCTTACCTGGATACCTGACAGG  
TGAGATGGCTGTAGCAGACACTGGCAGTTCCCTGCCCACACACCTGTCCCTGTCCACAGC  
TGCACAAGGCAGCTCTGTGTGCAATTGCCAGCATCTGCTCCTCTGTTCTCAGGGAATCTT  
TGTTAGAAAAATGCTGCCATATTTGTTTCTCACCTATTAGTCTTGTCTCCCAGTCAAGAG  
AATAAATTTATGCAAGCAGAGATTGTACTTTACAGTATTTTGTCTTTGAGCTTGGCATT  
[T, G]  
GTTGCATTTGTAAAAATGTGGCATGGCTTCCTCATCCCCAATAGGAACTTTGCCAGCCC  
TTTTGTTCTCATGGAACCTTCTTTTTTGAAGAGACACCAAAGGAGTAAAAATACTGTGG  
AGGGAGCAACCCTCCTTTGCCATATGCTCTCATTGGGAGACATGTGGAGCAGTCTGAAGT  
CATTAGGCCACTCTCTGGGAGAGCACATCCTATGATGTTCTCCAGCCTAGCCCCCTTCC  
ACTGTGCTCAAGTCCAAGCTGACCAGCTTCTGACCACAGTGTAACAAAGATGATTGTC

69134

CTGTGTGCAATTGCCAGCATCTGCTCCTCTGTTCTCAGGGAATCTTTGTTAGAAAAATGC  
TGCCATATTTGTTTCTCACCTATTAGTCTTGTCTCCAGTCAAGAGAATAAATTTATGCA  
AGCAGAGATTGTACTTTACAGTATTTTGTCTTTGAGCTTGGCATTAGGTTGCATTGTAA  
AAATGTGGCATGGCTTCCTCATCCCCAATAGGAACTTTGCCAGCCCTTTTGTTCATG  
GAACTTCCTTTTTTGAAGAGACACCAAAGGAGTAAAAATACTGTGGAGGGAGCAACCCT  
[C, T]  
CTTTGCCATATGCTCTCATTGGGAGACATGTGGAGCAGTCTGAAGTCATTTAGGCCACTC  
TCTGGGAGAGCACATCCTATGATGTTCTCCAGCCTAGCCCCCTTCCACTGTGCTCAAGTC  
CAAGCTGACCAGCTTCTGACCACAGTGTAACAAAGATGATTGTCAGTGGGCCCCAGAA  
TCCTATACCCAGA